Mind's eye: Saccade-related evoked potentials support visual encoding in humans

Gansheng Tan[1,2,3,4], Phillip Demarest[1,2,3,4], Yilin Li[1,2,3,4], Hohyun Cho[5], Haeorum Park[2,3,4], James R. Swift[2,3,4], Cory S. Inman[6], Joseph R. Manns[7], Stephan B. Hamann[7], Xiaoxuan Liu[1,2,3,4], Krista L. Wahlstrom[6], Ziwei Li[1,2,3,4], Martina K. Hollearn[6], Justin M. Campbell[6], Patrick E. Cettina[1,2,3], Siddharth S. Sivakumar[1,2,3], Eric C. Leuthardt[1,2,3,4], Jon T. Willie[1,2,3,4,8], Peter Brunner[1,2,3,4]

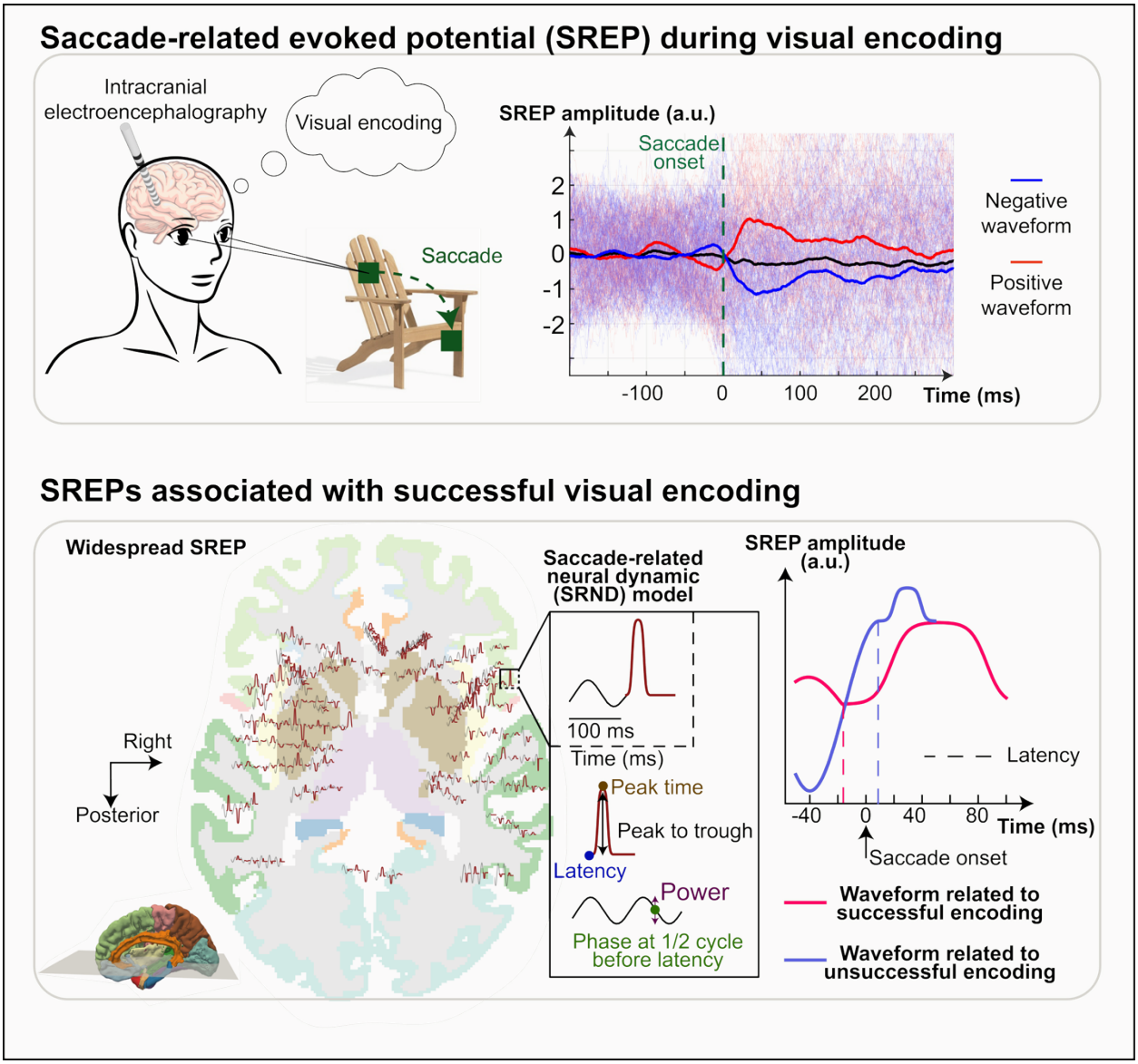
[1] Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, USA
[2] Department of Neurosurgery, Washington University School of Medicine in St. Louis, St. Louis, MO, USA
[3] Division of Neurotechnology, Washington University School of Medicine in St. Louis, St. Louis, MO, USA
[4] National Center for Adaptive Neurotechnologies, St. Louis, MO, USA
[5] Department of Neurosurgery, University of Alabama, Birmingham, AL, USA
[6] Department of Psychology, University of Utah, Salt Lake City, UT, USA
[7] Department of Psychology, Emory University, Atlanta, GA, USA
[8] Department of Neurosurgery, University of Texas at Austin, Austin, TX, USA

*communicating authors: Jon T. Willie and Peter Brunner

Highlights

- Widespread, consistent evoked potentials follow saccades (Saccade-related evoked potentials, SREPs) in human subjects with intracranial electrodes
- SREPs are not oculomuscular artifacts and carry no discernible information about saccade direction
- SREPs in the temporal lobes have earlier latencies and higher amplitudes than in other anatomical regions
- Earlier and high-amplitude SREPs across the cerebrum are associated with successful visual encoding

Graphical abstract



SUMMARY

In active vision, the brain receives and encodes discontinuous streams of visual information gated by saccadic eye movements. Saccadic modulation of neural activity is hypothesized to evolve to support perception and memory; however, it remains unclear whether this mechanism exist in humans, and potential functional roles have not been determined. We used eye tracking and intracranial local field potentials recorded from invasively monitored epilepsy patients when they performed visual encoding tasks, and observed consistent evoked potentials following saccades (saccade-related evoked potentials, SREPs) across the cerebrum. These SREPs were not attributable to ocular muscle activity or retinal input. Their magnitudes were not explained by spatial proximity to eye muscles or saccade eccentricity, and their polarity bore no relationship to saccade direction. Instead, the phase of pre-saccadic oscillation aligned with saccade timing and dissociated SREPs with positive polarity from those with negative polarity. Spatiotemporal profiling revealed that SREPs emerged earliest and with the greatest magnitude in the temporal lobes. We developed a saccade-related neural dynamic (SRND) model that characterized pre-saccadic oscillatory activity and SREP at each electrode contact location using finite features. Random forest models trained with these features achieved 62.6% balanced accuracy for predicting next-day recognition (long-term memory). Using the Shapley value, a framework for explaining machine learning models, we identified an SREP profile, characterized by earlier latency and larger magnitude, which was associated with successful visual encoding. In contrast, predicting saccade direction using the same SRND model performed at chance level, indicating that the observed SREP is less likely a corollary discharge signaling saccadic motor copy. These findings demonstrate SREPs as a neural mechanism of saccadic modulation with a role in human visual encoding.

Introduction

Every visual scan shapes our memory. Humans actively explore the environment through rapid eye movements that alternate between saccades and fixations. Saccades rapidly move the fovea to a new location in the visual field, and in the subsequent fixation, visual information from this location is sent to the visual system[1]. The brain processes the dynamic inflow of visual information structured by eye movements, constructs meaningful representations, and forms memories.

Despite the important role of saccades in gating visual input, many efforts to understand neurophysiological processes in visual encoding overlook the saccade-related neural dynamics.

Saccades in primates evoke potentials in the medial temporal cortex, inferotemporal cortex, and thalamus[2,3]. The saccade-related evoked potential (SREP) provides compelling evidence for saccadic modulation of neural circuits, as it synchronizes single-unit activity and accompanies changes in neural oscillatory patterns[2–5]. Although a plausible explanation is that saccadic modulation is a response to a shift of the visual image on the retina, findings of SREP in darkness and during rapid eye movement sleep challenge this hypothesis and suggest additional functions beyond visual response[6,7]. Understanding the functional role of SREPs in visual encoding is critical, as it may elucidate neural mechanisms underlying active vision and inform neuromodulation strategies for improving memory. However, the existence and functional role of SREPs in humans remained largely unknown due to technological constraints in simultaneous intracranial electrophysiology and eye tracking during naturalistic, free-viewing behavior.

Early research on visual processing focused on passive processing, which presented stimuli while the gaze was held constant. Although this approach allows precise control of when and where visual stimuli are presented, it sacrifices ecological validity. Recent advances in optical eye-tracking technology allows tracking gaze locations in real time. Simultaneous recording of eye movement and neural activity allows researchers to align neural dynamics with eye movements and visual input, opening new avenues to study visual processing in a more natural setting[8]. Extending previous findings of saccadic modulation in non-\human primates to humans requires recording neuronal activities with millisecond temporal precision. Neuroimaging, such as functional magnetic resonance (MRI), generally lacks the necessary temporal resolution. Noninvasive methods, such as scalp electroencephalography, provide millisecond temporal resolution, but are limited in source localization. Source localization remains a challenging inverse problem that depends heavily on assumptions and parameter choices. Oculomotor signal generated by eye muscles can confound the analysis of saccadic modulation, as reported in human electrophysiology[9,10].

One hypothesis of the functional role of saccadic modulation posits it as a form of *corollary discharge*, a copy of the motor signal that prepares the brain for the sensory consequences of an impending movement[11]. In vision, the corollary discharge mechanism maintains visual continuity despite the disrupted visual input resulting from saccades[12]. The neural pathways underlying the

corollary discharge mechanism have been previously identified in nonhuman primates. The projections from the superior colliculus (SC) to the frontal eye field via the medial dorsal thalamic nucleus and from the SC to the middle temporal cortex through the inferior pulvinar carry the vectors of impending saccades and contribute to visual continuity[13]. A recent study observed saccadic modulation of neuronal activity exhibiting some properties of corollary discharge was observed in the human mesial temporal lobe, a region for high-level visual processing and memory formation[2]. This finding comports well with the broader active sensing framework, which posits that motor sampling, such as sniffing in rodents and saccades in primates, is not random but part of a coordinated strategy that optimizes sensory information processing[14–17] (**Figure 1A**). Although these findings point towards a role of saccadic modulation in visual encoding, direct evidence in humans has yet to be established.

We are interested in determining whether saccadic modulation serves higher-order cognitive functions such as visual encoding in addition to basic perception, with a focus on SREPs. Two principal methods for quantifying neural dynamics in electrophysiological signals are evoked potential (EP) and spectral analysis. EPs measured by intracranial electrodes capture the rapid extracellular voltage deflection associated with neuronal activities following saccades[2]. The spatiotemporal dynamics of EPs can coordinate spike timing, either by directly influencing extracellular voltage or by generating rhythmic activities at saccadic frequency, thereby facilitating information processing. Spectral analysis typically requires longer time series that span multiple cycles of the assumed oscillation to accurately estimate its periodic properties[18]. This field of investigation is also experiencing a doctrinal shift from emphasizing traditional power spectrum analyses toward more nuanced approaches that distinguish strictly defined rhythmic components from the aperiodic $1/f$ activity in the frequency domain, while the methodology and interpretations of these components are still under active investigation[19]. Therefore, we sought to investigate the most rudimentary and well-established form of saccadic modulation, SREP, and test our hypothesis that SREPs have a role in visual encoding.

This study investigates the existence, spatiotemporal dynamics, and cognitive relevance of SREP by leveraging a unique research window: patients with epilepsy who undergo stereoelectroencephalography (SEEG) monitoring for evaluation for neurosurgery. To record neural activity during active vision, we expanded a previously developed bedside brain-computer interface research platform from our lab, which enables synchronous stimulus presentation, intracranial local field potential (LFP) recording, and eye movement tracking[20]. This combined technological advance enabled us to examine neural activity that was time-locked with saccades.

We first confirmed the presence of SREPs in humans, and verified that they represent saccadic modulation rather than ocular muscle signals. We observed two waveform patterns of SREP and examined potential determinants of these patterns, such as saccade direction and oscillatory activity. We found that the instantaneous phase of the preceding oscillations modulates the polarity of SREPs. To identify the neural pathway underlying saccadic modulation through EPs, we characterized the spatiotemporal dynamics of SREPs. We found that earlier and greater SREPs were spatially clustered in the temporal lobe. Next, we tested our central hypothesis that SREPs play a functional role in visual encoding using random forest models. Our results confirmed that EP characteristics predicted successful visual encoding. Building on this, we used Shapley value analysis to interpret the mapping between EP characteristics and the visual encoding learned by the models[21]. In general, earlier saccadic modulation with a greater magnitude and a later peak time predicted successful visual encoding. Finally, to assess whether these SREPs reflect a narrowly defined corollary discharge signal, we evaluated whether they encoded saccade direction[11]. We confirmed that most, if not all SREPs did not encode either the current or subsequent saccade direction. These results suggest that SREPs support visual encoding in the human brain, as a mechanism of saccadic modulation, separate from classical corollary discharge.

Results

Memory performance and eye movement characteristics during visual encoding

Eight subjects participated in visual encoding tasks that presented images, including objects and scenes for encoding. (**Figure 1B**, see Appendix A Study protocol for details). Subjects freely moved their eyes during encoding, and their recognition memory was assessed the next day. We tracked gaze coordinates and recorded intracranial electroencephalogram (iEEG) data during the tasks (**Figure 1E**). Data summary and patient characteristics are summarized in **Supplementary Tables 1–2**.

Across the eight subjects, a total of 1629 contacts spanning 129 electrode shanks were implanted, with an average of 204 ± 30 (mean ± SD) contacts on 16 ± 2 shanks per subject (**Figure 1C**). We identified and excluded electrodes exhibiting epileptiform activity or artifactual signals (e.g., due to poor contact) before re-referencing (**Supplementary Table 3**), resulting in an average of 201 ± 29 (mean ± SD) valid contacts per subject. Additionally for 6 subjects, two Behnke Fried Depth electrodes (AD-TECH), each with 8-contact micro recording, were implanted.

The distribution of electrode contacts across cortical and subcortical regions is shown in **Figure 1D**. When electrode contacts (~2 mm^3 volume) overlapped multiple brain regions, their contributions were proportionally allocated among these regions. Left white matter and left temporal cortex have a higher numbers of sampling electrode contacts (446 and 262, respectively) than other regions. The sampling density, defined as the ratio of the number of electrode contacts per unit anatomical volume, was greatest in the amygdala (13 contacts/cm^3), followed by the hippocampus (7 contacts/cm^3), and the left temporal cortex (4 contacts/cm^3, **Figure 1 – Figure supplement 2A**).

Depending on the patient's duration of stay in the epilepsy monitoring unit, three subjects completed two encoding tasks on two different days, one subject completed three encoding tasks on three different days, and each of the other four subjects completed one encoding task (**Supplementary Table 1**). This study is part of a clinical trial investigating amygdala stimulation in memory (NCT05065450) where amygdala stimulation was delivered in half of the encoding trials. To study the saccadic modulation in the normative visual encoding process, we only include data in valid encoding trials for subsequent analysis. A valid visual encoding trial had three criteria: (1) the subject pressed one of the response buttons once, (2) subjects maintained their gaze on the monitor for more than half of the trial duration as determined from eye-tracking data, and (3) no brain stimulation was delivered during the trial. The valid data includes a total of 1130 valid visual encoding trials and 7292 valid saccades. The data included for each subject are summarized in **Supplementary Table 1**.

Subjects remembered seen images at rates greater than chance (i.e., 50%), with a correct recall rate of 75% $\pm$ 16% (SD, standard deviation) and an overall correct recognition rate of 72% $\pm$ 8% (SD) (**Figure 1F, Figure 1 – Figure supplement 1F, Supplementary Table 3**). The mean fixation duration was 182.97 ms (SD = 265.92 ms); and the mean saccade duration was 47.79 ms (SD = 14.93 ms, **Figure 1G**). The saccade directions were dispersed, with a relatively higher probability of horizontal and vertical eye movements (35.54% and 26.45%, respectively, **Figure 1G**). The horizontal and vertical eye movements are saccades with directions falling within a $\pm 20°$ range centered on the horizontal axis or the vertical axis. To ensure no latent effects of electrical stimulation of the amygdala, we compared saccade behavior and recognition memory performance for trials with and without a preceding amygdala stimulation trial. The effect size (e.g., Cohen's d) for duration differences between trials with and without a preceding amygdala stimulation trial was -0.015 for saccades and -0.024 for fixation (**Figure 1 – Figure supplement 1C–D**). The circular distribution of saccade direction did not differ significantly between trials

following a cortical stimulation trial (N_1 = 3657) and those not following amygdala stimulation (N_2 = 3635, Watson–Williams test, p=0.879, F = 0.023, **Figure 1 – Figure supplement 1A–B**). The correct recall rate for trials that followed an amygdala stimulation trial, and those that did not, were both 75% (**Figure 1 – Figure supplement 1E)**. These results suggest that amygdala stimulation has a trivial impact of on subsequent visual encoding and saccades.

Consistent saccade-related evoked potential (SREP)

We first investigated whether saccadic modulation manifests as evoked potentials (EPs) in humans, extending observations previously documented in nonhuman primate studies. We excluded electrodes containing artifactual signals and used small Laplacian re-referencing to capture local electrical potential (see **Star Method**). We computed the inter-trial Pearson correlation coefficient on normalized signals to quantify the consistency of EPs following saccades[22]. We focused on saccades occurring during image presentation but excluded saccades occurring 200 ms before and after the image onset to isolate SREP. We expected the inter-trial correlation to be higher after saccades compared to the pre-saccade fixation (hereafter baseline). To our surprise, the distribution of inter-trial correlation was bimodal compared to baseline (**Figure 2A, Figure 2 – Figure supplement 2A**). Kolmogorov-Smirnov tests showed that the inter-trial correlation distribution following saccades significantly deviated from those during baseline (median p-value $<1 \times 10^{-20}$, significance threshold after correcting for multiple comparisons = 1.9×10^{-5}, 98% of electrode contacts showed significant deviation from baseline, **Figure 2 – Figure supplement 2B**).

We quantified the deviation from baseline (i.e., unimodal normal) distribution toward a bimodal distribution using the absolute area under the curve difference between the empirical cumulative distribution of intertrial correlation (**Figure 2 – Figure supplement 2A**). The deviation effect size corresponding to the adjusted significance threshold was 0.017, whereas the 2.5–97.5 percentile of the deviation effect size distribution was 0.011 to 0.228. These results suggest the presence of two clusters of EP waveforms following saccades. We identified these two clusters by applying agglomerative hierarchical clustering to the intertrial correlation matrix and partitioning the resulting dendrogram at the first bifurcation (**Figure 2B**). In case that a cluster was disproportionately small (less than 25% of the size of the other cluster), we removed this cluster and repeated the process. This occurred in 5 out of 2604 electrode contacts (0.2%), suggesting the intertrial correlation distributions for most electrode contacts were bimodal.

To quantify the EP waveform consistency after saccades compared to baseline, we computed the area under the receiver operating characteristic (ROC) curve for each cluster[22] (**Figure 2C**). The averaged EPs for the two clusters for a middle temporal lobe contact are shown in **Figure 2D**, where they exhibit opposing polarity and phase of pre-saccadic oscillatory activities. As a control, we repeated this process for baseline LFP to adjust for hierarchical clustering-induced area under the curve (AUC) inflation (**Figure 2E**). The AUCs following saccade were significantly higher than control AUCs, indicating that EP following saccade consistently have one of the two opposing waveforms (Cohen's d = 2.651, two-tailed paired t-test, n = 2604, t statistic = 81.13, $p < 0.001$, **Figure 2F**).

We quantified the response ratio at anatomical regions, defined as the proportion of electrode contacts with AUC in the peri-saccade condition above the mean + 2 SD of control AUCs (**Figure 2G**. The response ratios for all the cortical regions except the right motor cortex (37%) were above 50%. Fusiform gyrus (left: 92.0% right: 97.0%), temporal (left: 91.0%, right: 96.7%), and parietal lobes (left: 0.86%, right:100.0%) were the three regions with the highest response ratios. To ensure an objective assessment of saccade-evoked consistency across regions, we varied the responsiveness thresholds from mean to mean + 4 SD of control AUC. By integrating the response ratio over this range (i.e., computing AUC), we observed a similar spatial pattern of response ratio (**Figure 2 – Figure Supplement 1**). We ranked brain areas based on the area under the receiver operating characteristic (ROC) curve for peri-saccadic intertrial correlations, and their difference from control AUCs (**Figure 2 – Figure supplement 2C–E**). The AUC differences from control were generally higher in the temporal lobe (except the medial temporal lobe) and parietal lobe, followed by the frontal lobe, and subcortical regions. The electrode contact-level spatial pattern of AUC difference from control was consistent with previous findings (**Figure 2 – Figure Supplement 3A–E**). For LFPs recorded from microwires, peri-saccadic average AUCs were higher compared to controls, but were less prominent (**Figure 2 – Figure Supplement 3F–G**, Cohen's d = 0.227, two-tailed paired t-test, n = 115, t statistic = 2.088, $p < 0.039$).

These data demonstrate the presence of consistent SREPs of two waveforms in humans during active vision.

SREPs in iEEG reflect neural activity rather than ocular-muscle electromyogram artifacts

Although intracranial recordings are minimally affected by ocular-muscle electromyogram (EMG), a previous study showed that subdural electrocorticography grids can still capture some ocular-

muscle EMG signals[10]. If opposite saccade directions engage different pairs of extraocular muscles that produce electromyogenic fields of reversed polarity relative to the head, ocular-muscle EMG offers a straightforward explanation for the observed opposing SREP polarities. We tested two predictions that would hold if the observed SREPs reflect Ocular-muscle EMG signals: (1) their peak amplitude would decrease with increasing distance between the electrode contacts and the myogenic dipole sources generated by extraocular muscle activation, and (2) their peak amplitude would increase with greater saccade eccentricity. This theory also predict an association between saccade direction and SREP polarity, which we later show to be false.

We normalized SREP amplitude using a z-score based on the baseline and extracted the peak amplitude based on the averaged SREP. For SREPs of the same polarity, we used the MATLAB *findchangepts* function to identify two change points in the derivative of SREPs averaged across trials (**Figure 3A**). We defined latency as the first change point, and peak time as the time point within a 10-ms window around the second change point at which the z-scored SREP amplitude was maximal for positive clusters (or minimal for negative clusters). Similarly, we defined trough time as the time point within a 10-ms window around latency at which the z-scored SREP amplitude was minimal for positive clusters (or maximal for negative clusters). We approximate the distance between electrode contacts and the myogenic dipole source generated by extraocular muscle activation as the Euclidean distance between each electrode contact and the nearest eye globe center identified from the pre-operative T1-weighted MRI. Across all intracranial electrodes, the estimated slope between distance from the eyes and peak amplitude of averaged SREPs was trivial compared to the estimated intercept (**Figure 3D**, estimated slope = -0.002, estimated intercept = 1.078, explained variance = 3.8%, n = 2573 for positive polarity; estimated slope = 0.002, estimated intercept = -1.057, explained variance = 3.5%, n = 2575 for negative polarity, linear regression)**.**

To assess if peak amplitude increased with greater saccade eccentricity, we calculated the single-trial SREP peak amplitude by finding the amplitude peaks around the corresponding cluster-level peak times. As a control, we quantified the alignment in waveform between single-trial SREPs and the corresponding cluster-level averaged SREPs using Spearman's correlation. The distribution of alignment was skewed towards positive values with an area under the kernel density curve of 0.812 for correlations greater than 0 (**Figure 3B–C**). This indicates a general alignment between single-trial and averaged SREP waveforms. For trial-level analyses, we included trials with similarity above the median (i.e., Spearman's correlation $\geq$ 0.413). We calculated saccade eccentricity based on the eye and gaze location relative to the screen, derived

from eye-tracking data. We found that saccade eccentricity explained 0.0% of the variance in peak amplitude for both polarities (**Figure 3E,** |estimated coefficient| < 0.001, p < 0.001, n = 250531 for positive polarity; |estimated coefficient| < 0.001, p = 0.156, n = 250614 for negative polarity; linear regression with t-test**)**. As a positive control, we extracted electrooculogram (EOG) signals from fronto-polar EEG scalp electrodes (Fp1 and Fp2) using bipolar re-referencing, and computed the peak amplitude of EPs in EOG signals using the same methods. We found a greater correlation between saccade eccentricity and EP peak amplitude in EOG signals. Specifically, EP peak amplitude in EOG signals correlated positively with saccade eccentricity for positive polarity and negatively for negative polarity (**Figure 3F,** explained variance = 7.0%, estimated coefficient = 0.133, p < 0.001, n = 1100 for positive polarity; explained variance = 6.7%, estimated coefficient = -0.130, p < 0.001, n = 1126 for negative polarity; linear regression with t-test). At the electrode contact-level, the variance explained by saccade eccentricity was significantly greater in the ocular-muscular signals captured in fronto-polar scalp EEG electrodes compared to intracranial recordings (**Figure 3G**, t = 3.154, p = 0.012, Cohen's d = 3.193, $n_{EOG} = 10$, $n_{EOG} = 2367$, Welch's t-test).

To directly compare with prior work[10], we computed peak-to-trough amplitude as an complementary measure of SREP magnitude. Kovach et al. demonstrated a decrease in peak-to-trough amplitude of SREPs recorded in electrocorticography (ECoG) with increasing distance from the eyes[10]. On the contrary, both distance from the eyes and saccade eccentricity explained minimal variance in peak-to-trough amplitude across SEEG electrodes in our study. The distance explained 0.9% and 0.7% variance in peak-to-trough amplitude of averaged EPs for positive and negative clusters, respectively, across all electrode contacts (**Figure 3 – Figure Supplement 1A,** estimated slope = -0.001, estimated intercept = 1.311, n = 2571 for positive polarity; estimated slope = 0.001, estimated intercept = -1.287, n = 2570 for negative polarity; linear regression, data was omitted when we failed to detect trough). For both polarities, saccade eccentricity explained 0.0% variance in peak-to-trough amplitude (**Figure 3 – Figure Supplement 1B,** |estimated slope| < 0.001, p = 0.020, n = 250531 for positive polarity; |estimated coefficient| < 0.001, p = 0.079, n = 250614 for negative polarity; linear regression with t-test**)**. As a positive control, we examined the relationship between saccade eccentricity and peak-to-trough amplitude of SREPs in EOG signals. We found that EP peak-to-trough amplitude in EOG signals correlated positively with saccade eccentricity for positive polarity and negatively for negative polarity (**Figure 3 – Figure Supplement 1C**, explained variance = 2.7%, estimated coefficient = 0.108, p < 0.001, n = 1100 for positive polarity; explained variance = 8.2%, estimated slope = -0.171, p < 0.001, n = 1126 for negative polarity; linear regression with t-test).

Saccade of longer duration generally involves more sustained or stronger muscle activity. In this study, we found that saccade duration positively correlated with saccade eccentricity but not with absolute peak amplitudes of the SREPs (**Figure 3 – Figure Supplement 1D**, Pearson correlation coefficient = 0.056, $p < 0.001$ between saccade duration and saccade eccentricity, Pearson correlation coefficient = -0.002, $p=0.080$ between saccade duration and absolute peak amplitudes, permutation test). These results are consistent with the conclusion that intracranial SREPs were not related to oculomotor activity.

We next examined whether the SREPs represented the change in visual input resulting from saccades. If this is the case, we would not expect consistent LFP waveforms following saccades that occurred in the absence of meaningful visual changes, that is, during the intervals between images when a blank screen containing a fixation cross on a uniform black background is presented. We computed the SREP consistency during fixation cross presentation periods using the same method described above. Despite the absence of changes in visual inputs, SREP consistency during fixation cross presentation was significantly higher than baseline (**Figure 3H**, paired t-test, $p<0.001$, Cohen's $d = 2.631$, $n = 2604$). Across electrode contacts, SREP consistency during image presentation was positively correlated with that during fixation periods (**Figure 3I**, Pearson correlation coefficient = 0.752, $p < 0.001$, based on t-distribution). This correlation pointed to a regional specificity in how saccades modulate LFPs independently of concurrent visual input. We calculated the ratio of electrode contacts showing consistent SREPs across anatomical regions during fixation cross presentation and found the spatial pattern was similar to that during image presentation (**Figure 3 – Figure Supplement 1E**, **Figure 2G**).

These results demonstrate that SREPs are not attributable to changes in visual input nor ocular-muscular EMG signals, which motivates our subsequent analysis of their spatiotemporal dynamics.

Systematic relationship between the phase of preceding oscillatory activity and the polarity of SREPs

In theory, saccade direction influences the polarity of voltage deflections in EOG signals due to volume-conducted oculomotor artifacts. We observed a unimodal distribution of saccade directions associated with monopolar SREPs derived from EOG signals (**Figure 4 – Figure Supplement 1A**). To quantify this directional bias, we computed Rayleigh statistics based on the saccade direction distributions associated with positive and negative SREPs, respectively. To control for the influence of sample size, we resampled the same number of saccade directions for positive, negative, and combined polarities. The polarity-specific Rayleigh statistics for EOG

signals were significantly higher than those computed across all saccade directions, indicating directional clustering (**Figure 4 – Figure Supplement 1B,** t=5.950, p < 0.001, Cohen's d = 1.881, n = 15 for positive polarity; t = 4.553, p = 0.001, Cohen's d = 1.440, n = 18 for negative polarity; paired t-test). In contrast, there was no evidence of directional clustering for saccade directions associated with monopolar SREPs that were derived from LFPs measured by SEEG electrodes. For both polarities, Rayleigh statistic differences (monopolar – combined) were centered at zero (**Figure 4 – Figure Supplement 1C,** t=-0.953 p = 0.341, Cohen's d = -0.022, n = 2068 for positive polarity; t = -1.028, p = 0.304, Cohen's d = -0.023, n = 2058 for negative polarity; two-tailed t-test).

We further tested whether the SREP polarity was associated with different saccade direction distributions by comparing the difference between the saccade directions associated with positive and negative SREPs using Mardia-Watson-Wheeler tests. The circular distribution of saccade direction for EOG signals was significantly different between positive and negative polarity (p < 0.001 for all subjects, 95% CI for W statistics = [33.223, 46.050], Mardia-Watson-Wheeler test). By contrast, there was no significant difference in saccade direction between polarities in SREPs derived from LFPs (95% CI for W statistics = [1.926, 2.107], the corresponding 95% CI for p-value is [0.486, 0.512], Mardia-Watson-Wheeler test).

The absence of directional bias associated with positive or negative SREPs recorded in intracranial electrodes prompted us to consider whether pre-saccadic neural dynamics affected SREP polarity. Given the ubiquitous presence of oscillatory activities in LFP and previous evidence showing a phase-locking relationship between saccade timing and alpha oscillations, we hypothesized that the phase of preceding oscillatory activity modulates SREP polarity[23]. To test this, we identified oscillatory activities at each electrode based on the estimated power spectral density. We identified the dominant oscillation frequency by fitting a one-over-frequency ($1/f$) curve. We defined the dominant oscillation frequency as the frequency at which the residual of the power spectral density from the $1/f$ fit was largest (**Figure 4A**). Across electrode contacts, the distribution of dominant oscillation frequency centered at 11 Hz (**Figure 4 – Figure Supplement 1D**). We quantified the power of these oscillatory activities as the deviation between the power spectral density and the $1/f$ fit at the dominant frequency. **Figure 4 – Figure Supplement 1E** shows the distribution of the power of oscillatory activities. The mean oscillation power was 5.9 $\mu V^2/Hz$, suggesting non-trivial oscillatory activities.

We estimated the instantaneous phase and power by applying the Hilbert transform to padded signal segments preceding SREP latency, narrowband-passed at ± 2 Hz around the dominant oscillation frequency (**Figure 4A**). This preprocessing avoided the potential influence of

SREP in phase estimation. Using SREP latency as a temporal reference, we defined five time points of interest: 0.5, 1, 1.5, 2, and 3 oscillatory cycles (OCs) before SREP latency. To test if the phase of pre-saccadic oscillatory activity modulates the polarity of SREPs, we first examined the distribution of oscillation phases preceding positive and negative SREPs for each electrode contact. We estimated circular probability density functions for the instantaneous phase values associated with each polarity at the five time points. We quantified phase separation as the area under the absolute difference between the two probability density curves. The phase distribution of preceding oscillatory activity recorded in an electrode contact displayed greater phase separation as time approached SREP latency (**Figure 4B**). To assess whether the phases preceding SREP were different for positive and negative SREPs, we generated surrogate (i.e., null) distributions by randomly permuting the polarity labels. For each electrode contact, we computed the phase separation metric on permuted data using the same procedure applied to the observed data. Across electrode contacts, we found significant phase separation between polarities at the five time points of interest, compared to the null distributions (**Figure 4C,** $p < 0.001$ for all time points, paired Wilcoxon signed-rank, $n = 1970$ pairs). The degree of separation increased from 3 to 0.5 OCs prior to EP latency, with corresponding rank-biserial correlation coefficients of 0.204, 0.581, 0.790, 0.901, and 1.000.

To determine if pre-saccadic oscillation phase-SREP polarity relationship differs across anatomical regions, we computed the circular mean phase for each polarity and electrode contact by averaging across trials. Here, trial refers to a single SREP, distinct from a trial in the behavioral task, which is defined as one image presentation. The distribution of mean phase values across electrode contacts was separated by polarity (**Figure 4D**). At 1 OC before SREP latency, the mean phase values associated with positive SREPs were near $\pm\pi$, corresponding to the trough of the oscillation, whereas the mean phase values associated with negative SREPs were near 0, corresponding to the oscillation peak. This separation of the mean phases between polarities was more pronounced as time approached SREP latency (**Figure 4D**). These results suggest a systematic relationship between SREP polarity and the phase of pre-saccadic oscillation.

Oscillation power and phase can adjusting neuronal sensitivity to sensory information based on the active sensing theory[24]. To understand how SREP interacted with oscillatory activities, we first computed the phase lag index across trials, separately for those associated with positive and negative SREPs. For both polarities, the phase lag index continuously increased from 2 OCs before latency and started to decrease after peak time (**Figure 4E**). The effect size (Cohen's d) between phase lag index at 2 OC before latency and subsequent time points (from 1

OC before latency to 1 OC after peak time) were 1.214, 1.625, 1.739, 1.704, and 1.427. These findings indicate that SREPs are preceded by a gradual phase alignment of oscillatory activities, suggesting they are embedded within a broader window of saccade-timed neural coordination.

Phase alignment reflects the temporal coordination of neural ensembles. We next asked whether it was accompanied by changes in oscillatory power, which represents synchronization of local neural ensembles. We averaged oscillation power across SREP events and normalized it to power at two OCs before the SREP latency for each electrode contact, enabling comparison across electrode contacts and subjects. Oscillatory power increased from 2 OCs before the SREP latency, and remained elevated (**Figure 4 – Figure Supplement 1F**). We further evaluated the correlation between phase alignment and power increase. We found the strongest positive correlation at SREP latency (**Figure 4F**, $R^2 = 0.274$). The correlation was weaker at earlier and later time points, with R^2 = 0.022 at 1 OC before SREP latency, R^2 = 0.118 at 0.5 OCs before SREP latency, R^2= 0.198 at the SREP peak, followed by a decline to R^2 = 0.046 at 0.5 OCs after the SREP peak. The differences in the temporal profiles of phase alignment and oscillation power suggest that they may reflect distinct neural mechanisms. The early increase in oscillation power may reflect the preparation for effective information processing, whereas subsequent phase alignment relative to the times of saccades may facilitate information integration across regions[25].

These results show an interplay between intrinsic rhythmic activity and evoked potential, and that the pre-saccadic oscillation phase dissociates SREP waveforms.

Spatiotemporal profiling reveals early SREPs in the temporal lobe

If SREPs play a role in visual encoding, their characteristics should vary across task trials and predict trial-by-trial behavioral performance, namely recall. To assess whether these SREPs reflect dynamic neural processes or represent stereotyped, location-specific responses, we analyzed both intra-contact and inter-contact variability in SREP latency and amplitude. Consistent SREP latencies across trials at each recording site would imply stereotyped responses, whereas variability within electrode contacts would indicate modulation by internal brain states. To ensure that the estimation of intra-contact and inter-contact variability was based on matched sample sizes, we performed resampling procedures. For intra-contact variability, we resampled single-trial SREP latencies within each contact to match the median trial count across all electrode contacts. Similarly, for inter-contact comparisons, we resampled the set of electrode contacts to equalize the number of latency estimates per contact. The average intra-contact standard deviation of SREP latencies was 5.3 ms for both polarities (**Figure 5A**), which was non-trivial given the time interval between peak time and latency (mean = 38.6 ms, SD = 12.3 ms). The

comparison between intra-contact variability and inter-contact variability of SREP latency showed that the inter-contact variability was significantly greater, with a standard deviation of 7.1 ms and 6.6 ms for negative and positive SREPs, respectively (**Figure 5A**, N = 786, W statistics = 24141, $p < 0.001$, rank-biserial correlation -0.766 for negative polarity; N = 797, W statistics = 34836, $p < 0.001$, rank-biserial correlation -0.669 for positive polarity; Wilcoxon signed-rank test). This finding suggests that although variability in SREP latency existed within regions, more variability arises across regions, suggesting a spatial gradient of SREP latency across brain regions.

To verify the spatial specificity of SREP latency, we computed the latency of averaged SREPs separately for positive and negative polarities at each electrode contact and assessed their correlation across electrode contacts. We found a significantly positive correlation (Pearson $r = 0.436$, $p < 0.001$; **Figure 5 – Supplement 1A–B**), suggesting that each anatomical site was associated with a characteristic SREP latency, regardless of polarity. We also found a positive correlation between the difference in latency derived from averaged SREPs across electrode contacts and the effect size derived from single-trial SREP latencies (**Figure 5 – Supplement 1C**, Pearson $r = 0.436$, $p < 0.001$, t-test). Together, these results indicate that SREP latency across anatomical regions captures the spatiotemporal dynamics of SREP.

Spatiotemporal profiling revealed that electrode contacts with earlier SREPs were clustered in the temporal lobe (**Figure 5B**). To characterize this pattern anatomically, we identified electrode contacts with SREP latency in the 0–15th percentile range and applied 3-dimensional (3D) Gaussian Mixture Models (GMMs) to their spatial locations. This analysis revealed two predominant clusters located in each temporal lobe. Ee computed the average SREP latency across anatomical regions, and confirmed that the SREPs in bilateral temporal lobes had earlier latencies than other regions. This may suggests a sequence of the saccadic modulation pathway originate in temporal lobes (**Figure 5C**). SREPs of both polarities displayed similar latency patterns across regions (**Figure 5 – Figure supplement 1E–F**).

To examine whether early SREPs were more spatially localized, we stratified electrode contacts by SREP latency percentiles and quantified spatial spread as the Euclidean norm of the standard deviations of contact coordinates along three orthogonal axes (Left-right, posterior-anterior, and inferior-superior). We found that spatial spread increased progressively from early to late SREPs (**Figure 5D, Figure 5 – Figure supplement 1D**). **Figure 5D** also shows regions where SREP latencies were within three representative latency ranges: early (0–15th percentile), mid (30–45th percentile), and late (60–75th percentile). Early SREPs were overwhelmingly observed in the bilateral temporal lobes and insula, mid-latency SREPs were predominantly in

parietal, subcortical regions (basal ganglia and thalamus), and medial temporal (amygdala and hippocampus) region, and late SREPs were distributed more widely across anatomical regions. To complement the across-region comparison, we assessed the changes in the proportion of electrodes whose SREP latency fell within a given latency percentile across bins for each region. Specifically, we computed Spearman's correlation between the latency percentile bin and the proportion of electrodes whose SREP latency was in that bin. We found strong negative correlationd in the middle (Spearman's $\rho = -0.97$, corrected $p < 0.001$, number of comparisons = 20), inferior ($\rho = -0.74$, corrected $p = 0.455$), and superior ($\rho = -0.98$, corrected $p < 0.001$) temporal regions, indicating a higher fraction of early SREPs with a progressive decline across later bins. In contrast, increasing number of electrode contacts in the orbital ($\rho = 0.82$, Bonferroni corrected asymptotic $p = 0.136$) and inferior ($\rho = 0.89$, corrected $p = 0.025$) frontal regions showed later SREPs. Together, these results support saccadic modulation that may first occurs in the temporal lobe and subsequently recruits broader sites.

Intra-contact variability of SREP amplitude was higher than inter-contact SREP amplitude (**Figure 5 – Figure supplement 2A**). The non-trivial range of intra-contact variability suggests that SREP amplitude might represent the evolving neural dynamics associated with each saccade. We examined the peak amplitude distribution of averaged SREPs across electrode contacts, and found that it was symmetric for positive and negative polarities (**Figure 5 – Figure supplement 2B**). Across electrode contacts, the peak amplitudes of positive and negative averaged SREPs were negatively correlated (**Figure 5 – Figure supplement 2C**, Pearson's correlation coefficient = -0.365, $p < 0.001$). These findings indicate that peak amplitudes of positive and negative SREPs are dependent, suggesting region-specific amplitude scaling. For a pair of electrode contacts, the effect size between distributions of single-trial SREP amplitude was positively correlated with the difference in peak amplitude in the averaged SREP (**Figure 5 – Figure supplement 2D**, Pearson's correlation coefficient = 0.510 for negative SREPs, Pearson's correlation coefficient = 0.479 for positive SREPs, $p < 0.001$ for both polarities, exact test for Pearson's correlation). These results suggest anatomical specificity, motivating the analysis of the spatial gradient of SREP peak amplitudes.

We employed the same approach for quantifying the spatial pattern of SREP peak amplitude as for latency. Consistently, SREPs with higher peak amplitude were clustered in the bilateral temporal lobes (**Figure 5F**). The absolute peak amplitudes averaged across electrode contacts localized within an anatomical region were highest in the bilateral temporal lobe (**Figure 5 – Figure supplement 2G**). Similarly, we stratified electrodes by percentile bins of absolute peak

amplitude and measured spatial spread, which decreased monotonically with increasing peak amplitude percentile bins (**Figure 5G** and **Figure 5 – Figure supplement 2E**). These results indicate that stronger SREPs formed localized spatial clusters, whereas those with smaller peak amplitudes were more spatially dispersed. We performed the same region-wise correlation analysis between peak amplitude percentile bin and the proportion of electrodes whose SREP peak amplitude fell in that bin. We found that anterior (**Figure 5 – Figure supplement 2F,** $\rho = 0.95$, Bonferroni corrected asymptotic $p = 0.002$, number of comparison=20), inferior ($\rho = 0.95$, corrected $p = 0.001$), and middle ($\rho = 0.93$, corrected $p = 0.005$) temporal regions had a lower proportion of electrode contacts showing low absolute-amplitude SREPs with and a higher proportion of electrode contacts showing SREPs with high absolute-amplitude. The majority electrode contacts in white matter ($\rho = -1$, corrected $p < 0.001$) showed SREPs with low absolute-amplitude, suggesting that SREP in white matter may represent attenuated volume-conducted electrical field from synchronous neural activities.

These analyses revealed that, despite intertrial variability, particularly in peak amplitude, SREP followed a systematic anatomical progression. Our results indicate leading emergence of saccadic modulation from temporal lobe neocortex subsequently propagate to broader cortical and subcortical networks.

SREPs correlate with visual encoding outcome

Built upon the observed trial-by-trial variability in SREP waveforms, we next asked if this variance explained visual encoding outcome (i.e., remembered versus forgotten). Because each SREP waveform is a high-dimensional time series, learning a mapping from the full waveform to memory outcome would be computationally inefficient and difficult to interpret biologically. To obtain a physiologically meaningful representation, we leveraged our earlier findings about the interplay between pre-saccadic oscillation and the subsequent SREP waveform. We therefore defined a saccade-related neural dynamic (SRND) model for each saccade event and electrode contact, decomposing the neural dynamics into two components: (1) pre-saccadic oscillatory activity, and (2) the SREP waveform that follows. Specifically, we parameterized these two components: pre-saccadic oscillatory activities and the subsequent SREP waveform. We described these two components using five parameters: SREP latency, SREP peak amplitude, SREP peak time, and the power and phase of the ongoing oscillation at half oscillation cycle before the saccade onset (see **Star Methods**). We then applied machine learning to map SRND model parameters across anatomical regions to successful visual encoding. Successful visual encoding was defined by the subject's ability to recognize previously viewed images on the following day. **Figure 6A** shows

reconstructed preceding oscillation and SREPs across anatomical regions during a remembered versus a forgotten trial. As a control analysis, we verified that saccadic behaviors themselves did not predict visual encoding (**Figure 6B**). Saccade eccentricity and duration were comparable between saccades during remembered and forgotten trials ($n_{remembered}$ = 3913, $n_{forgotten}$ = 1227, t = -0.325, p = 0.745, Cohen's d = -0.011, t-test for saccade eccentricity; t = -1.237, p = 0.216, Cohen's d = -0.0039, t-test for saccade duration). To assess differences in saccade direction between remembered and forgotten trials, we estimated circular densities of saccade angles and quantified between-condition differences as the absolute area between the two densities over [–π, π]. By contrast, a permutation test, in which condition labels were shuffled, showed no significant difference in saccade direction between saccades in remembered and forgotten trials (p = 0.862).

To study the potential non-linear relationship between SRND parameters and visual encoding, we trained random forest models[26]. A random forest consists of an ensemble of decision trees, where each tree partitions the feature space through a sequence of split decisions. Splitting proceeds until a terminal leaf node is reached. Each leaf stores the empirical class distribution of the training samples that reached it, which defines the predicted probability for each class. For each tree, a new observation follows a path to a leaf node based on the decision rules. The aggregated probability across trees gives the final prediction. Training involves identifying subsets of features and split thresholds that maximize impurity reduction (Gini index) at each decision node[26]. In this study, the input features included the spatial location of electrodes and the five SRND parameters. To rigorously assess predictive validity, we split the data into training and testing sets at the trial level so that all saccades occurring in a single trial (e.g., encoding one image) and their related SREPs recorded in all electrode contacts were confined to either the training or the test set. Consequently, the model was never exposed to neural data from the same image during training and testing. This prevents the model from leveraging idiosyncratic properties of specific images, ensuring that the learned mapping reflects the relationship between SRND parameters and encoding success. To minimize subject-specific confounds while enabling learning mapping specific to spatial location, we added normally distributed spatial jitters to electrode contact location. The model architecture is shown in **Figure 6c**. Because the number of remembered and forgotten trials was imbalanced, we used balanced accuracy as our primary outcome to determine whether neural dynamics around saccades predicted visual encoding (**Supplementary Table 3**).

In 5-fold cross-validation (CV), the model achieved 55.7% balanced accuracy and 70.3% accuracy at the electrode contact level, exceeding all values in the null distribution generated by permutating encoding success labels (**Figure 6d, Figure 6 - Figure Supplement 1a**). This suggests that SREPs recorded in individual electrode contacts carried information predictive of visual encoding. As successful visual encoding likely depends on coordinated activity across neural networks, we evaluated the predicted probability at the saccade level. We used SRND parameters for all locations during a saccade as input to the model and obtained the average predicted probability. The observed ROC AUC on test data (0.662) exceeded all values from a null distribution generated by training and evaluating models on data with permuted labels (**Figure 6e**). The observed saccade-level balanced accuracy was 62.6%, exceeding all values in the null distribution. Moreover, the balanced accuracy at the electrode contact and saccade level for all CVs exceeded all values from the null distributions (**Figure 6 – Figure Supplement 1f**).

To assess whether consecutive SREPs in a trial reflected a stable neural state for visual encoding, we used the SRND parameters for all electrode contacts across all saccades during a trial as input to the model and computed the average predicted probability as the trial-level predicted probability of successful visual encoding. At the trial-level, the observed ROC AUC on test data was 0.678, and balanced accuracy was 63.1% (**Figure 6 - Figure Supplement 1b**). We confirmed that the average predicted probability for testing data was higher in remembered trials, than in forgotten trials (**Figure 6 – Figure Supplement 1c**, U statistic = 84652, $p < 0.001$, rank-biserial correlation = 0.333, $n_{remembered}$ = 635, $n_{forgotten}$ = 200, Mann–Whitney U test). Although both testing ROC AUC and balanced accuracy at the trial level exceeded all values from the null distributions generated by training and evaluating models on data with permuted labels, they were close to the saccade-level performance. In other words, aggregating SREP parameters across multiple saccades in a trial did not substantially improve predictive accuracy, suggesting that not all SREPs during a successful visual encoding trial reflected a neural state conducive to visual encoding.

To evaluate the possibility that SREPs reflected moment-to-moment neural states tied to individual saccades, we examined model predictions at the saccade level. Specifically, we computed the proportion of saccades the model classified as occurring in a remembered trial, thereafter termed encoding-conducive saccades. 75% of saccades in actual remembered trials were encoding-conducive saccades, and 49% of saccades in forgotten trials were encoding-conducive saccades (**Figure 6 – Figure Supplement 1D**). A chi-squared test of independence revealed a significant association between trial condition and predicted saccade labels ($\chi^2(1)$ =

217.89, $p < 0.001$). These results have two implications. First, remembered trials had a greater proportion of SREPs predicted to be associated with successful encoding, suggesting that the random forest model has learned SREP patterns conducive to visual encoding. Second, the proportion of SREPs predicted to be associated with successful encoding was below 100% in remembered trials, confirming that not every saccade in a remembered trial occurred in a neural state conducive to encoding. Conversely, in forgotten trials, not all saccades were predicted to be associated with unsuccessful encoding, implying that even when visual encoding ultimately failed, transient neural states supportive of encoding can still arise but may be insufficient or poorly timed to drive successful visual encoding.

To verify that the random forest model captured meaningful relationships without fitting noise (e.g., neither underfitting nor overfitting), we systematically evaluated how hyperparameters influenced model performance. We varied the number of trees, maximum tree depth, and minimum sample size per leaf. As expected, increasing tree depth and decreasing the minimum leaf size allowed the model to learn more complex relationships, leading to higher training accuracy until saturation (**Figure 6F**). Importantly, test accuracy rose in parallel with training accuracy and did not decline. We used the default parameters in the *imbalanced-learn* Python package (200 trees, maximum depth = 20, minimum leaf size = 1). Under this setting, the model did not overfit and achieved relatively higher training and testing accuracy compared to other hyperparameter configurations, as evidenced by the absence of divergence between training and testing accuracy within a region of hyperparameters surrounding the default setting.

These results establish SREPs as physiological correlates of a neural state that supports visual encoding.

Profile of SREPs that predict successful visual encoding

Our findings show that SRND parameters predicted successful visual encoding, suggesting that certain SREP profiles may provide a biomarker reflecting an optimal neural state that facilitates visual encoding. To describe this biomarker profile, we explained the learned mapping between SRND parameters and encoding outcomes using Shapley values[27]. As a first step, we examined whether the mapping learned by the random forest models was consistent across CVs by quantifying (1) the concordance of feature importance, (2) the similarity of split histograms, and (3) the consistency of model outputs across CVs. We compared these values against null distributions generated by permuting the successful encoding labels once to create a surrogate dataset, from which we trained random forest models using 5-fold CVs. We used Gini impurity reduction to quantify the feature importance and used Kendall's coefficient of concordance to

assess the concordance of feature importance across CVs. The concordance coefficient across CVs (Kendall's W = 0.966) for the actual data was higher than all values in the null distribution (mean Kendall's W = 0.439, SD = 0.110), suggesting that the models across CVs consistently identified the same features as important and weighted them similarly (**Figure 6G**).

Next, we examined whether the decision thresholds learned by the model were consistent across CVs. A split histogram, which represents the distribution of decision thresholds across features, captures the structure of these decision boundaries. We compared split histograms between pairs of CV iterations and quantified their similarity using the area under the absolute difference curve. The split histogram for models trained on actual data across CVs were more similar than the null distribution generated based on permuted data (U = 0.000, p < 0.001, Mann-Whitney U test, Cohen's d = -3.398, $n_{actual} = 10, n_{null} = 100$), suggesting that random forests trained on actual data exhibit consistent decision boundaries across CVs (**Figure 6H**).

To assess whether random forest models trained on actual data yield consistent predictions across CVs, we randomly sampled an observation from the test set in one of the CV iterations and used it as input to two models trained in different CV iterations. We repeated this procedure 200 times, both for models trained on actual data and on models trained on a dataset generated by permuting labels. We used Pearson's correlation coefficient between the predicted probabilities from a pair of models to quantify the output agreement between the two models. Models trained on actual data had higher output agreement across CVs than those trained from permuted data (**Figure 6I**, U = 40000, p < 0.001, Mann-Whitney U test, Cohen's d = 5.208, $n_{actual} = 200, n_{null} = 200$). We assessed output agreement based on the ratio of samples with different predictions. For a pair of models trained on actual data in different CV iterations, we used the same test data as input to both models and computed the proportion of samples with different predictions. this proportion was significantly lower than the null distribution derived from models trained with permuted data in cross-validations (**Figure 6 – Figure Supplement 1H**, U = 0, p < 0.001, Mann-Whitney U test, Cohen's d = -7.656, $n_{actual} = 200, n_{null} = 200$). Together, these results demonstrate that models across folds captured the same underlying mapping between SRND parameters and visual encoding, laying the groundwork for explanatory analyses to reveal the SREP profiles for successful visual encoding.

To determine how individual features contributed to the predicted probabilities, we calculated Shapley values for SRND parameters based on a trained random forest model. Shapley values quantify how much a given feature value increases or decreases the probability of successfully encoding of an image. The Shapley value is the average marginal contribution of

a feature value across all possible coalitions. This framework provides attribution of the probability of remembering at the level of individual observations and accounts for correlations between features. **Figure 7A** illustrates the distribution of Shapley values for SREPs recorded from a representative electrode contact, from which we can infer feature values associated with successful visual encoding. For example, smaller latency values were associated with higher Shapley values, indicating that earlier SREPs at this location increased the probability of successful remembering. The distribution of Shapley values for SRND parameters aligned with the patterns shown by partial dependence plots (**Figure 7B**). As another example, higher absolute peak-to-trough amplitudes were associated with larger Shapley values, which was reflected in both the feature value – Shapley value relationship in **Figure 7A** and the U-shaped partial dependence curve in **Figure 7B**.

To better illustrate how Shapley values represent the mapping learned by a random forest model, we systematically varied SRND parameters, including instantaneous phase and power at 0.5 OCs before SREP latency, peak-to-trough amplitude, and peak time. We used the feature combinations and the contact's spatial coordinates as inputs for the trained random forest model to obtain the predicted probability of successful visual encoding. **Figure 7C** shows the probability of successful encoding for reconstructed SREPs. We found that SREPs with early latency and high amplitude were associated with successful visual encoding, consistent with the feature value-Shapley value relationship shown in **Figure 7A**.

To better interpret the learned mapping between SRND parameters and visual encoding, we constructed a lower-dimensional representation of the random forest model. This representation consisted of a threshold and a direction for each feature at each spatial location, derived from Shapley values. Conceptually, this procedure mimics the decision process of a random forest and allows us to identify the range of SRND parameters conducive to visual encoding. For non-circular features (e.g., peak-to-trough amplitude), we ordered observations by feature value and identified a change point that minimized the within-segment variance of Shapley values, partitioning values into encoding-favorable versus unfavorable ranges (**Figure 7D**). For the circular feature (e.g., the instantaneous phase of preceding oscillations), we identified the angular window on the unit circle that minimized the variance of Shapley values inside versus outside the window (**Figure 7D**). The direction associated with successful encoding was then defined by comparing mean Shapley values between the two sides of the threshold (or inside vs. outside the window).

To confirm that the Shapley value-informed thresholds and directions explain visual encoding, we calculated a recall likelihood, a probability-like measure of successful encoding based on these thresholds and directions. For each observation, we computed the distance between the feature value and the threshold, scaled by the standard deviation of the feature values. To calculate the recall likelihood, we added or subtracted the scaled distance from the prior probability (e.g., 0.5), depending on the Shapley value-informed direction, and then clipped the result to the range [0, 1]. For the circular feature, we calculated recall likelihood in a similar way by adding or subtracting the shortest angular distance to the boundaries from the prior probability. Recall likelihood averaged across electrode contacts was significantly different between remembered and forgotten trials (**Figure 7D,** $N_{remember} = 3052$, $N_{forgotten} = 906$, U = 1.872×10^6, p < 0.001, Mann-Whitney U test, rank-biserial correlation coefficient = 0.354).

We confirmed that the Shapley value-informed thresholds and directions predicted successful visual encoding. **Figure 7E** shows the balanced accuracy of distinguishing between remembered and forgotten trials using recall. For each saccade, we computed the recall likelihood of every SRND parameters at every electrode contact and compared it against a baseline probability of 0.5. We generated the control thresholds and directions by randomly sampling thresholds from the observed feature values. For all features, the balanced accuracy was higher using recall likelihood based on actual Shapley value-informed threshold and direction ($N_{observation} = 2707$, $N_{control} = 5414$, p < 0.001, t test, Cohen's d ranged from 0.296 to 0.390). The balanced accuracy of separating successful from unsuccessful visual encoding using recall likelihood for all features was significantly higher than the control distribution ($N_{observation} = 2707$, $N_{control} = 5414$, U = 9.767×10^6, p < 0.001, Mann-Whitney U test, Cohen's d = 0.571). **Figure 7 – Figure Supplement 1A** shows the area under the receiver operating characteristic (ROC) curve for classifying whether a SREP recorded at an electrode contact during a saccade occurred in a remembered or forgotten trial, based on recall likelihoods. To generate a null distribution for comparison, we repeated this procedure using control thresholds and directions, which were randomly sampled from the observed distribution of feature values. The area under the curve (AUC) was significantly higher using recall likelihood based on actual Shapley value-informed threshold and direction ($N_{observation}$ = 2707, $N_{control}$ = 5414, p < 0.001, Mann-Whitney U test, U = 1.067×10^7, Cohen's d = 0.817), suggesting that Shapley-informed recall likelihoods captured meaningful structure distinguishing remembered from forgotten trials at the electrode contact level.

We next validate that Shapley value-informed thresholds and directions across electrode locations reflect neural states associated with successful encoding. Specifically, we calculated the

proportion of SRND parameters associated with successful visual encoding across electrode locations for each saccade, and constructed a receiver operating characteristic (ROC) curve by varying a proportion threshold. The averaged AUC across subjects was 84.9%, exceeding all values in the control distribution (**Figure 7F** , rank-biserial correlation = 1.000, Cohen's d = 6.789). The averaged balanced accuracy at the optimal proportion threshold across subjects was 77.9%, which also exceeded all values in the control distribution as well (rank-biserial correlation = 1.000, Cohen's d = 8.761). We generated the control distribution by randomly sampling thresholds from the feature values. Notably, Shapley values–informed thresholds and directions better predicted successful visual encoding at the saccade level than the random forest classifier (**Figure 6E**). This suggests that the low-dimensional threshold–direction representation preserves and better identifies the mapping between visual encoding and SREPs by extrapolating the complex, tree-structured decision rules of the random forest into a more generalized form that is less sensitive to noise. **Figure 7 – Figure Supplement 1B** shows the area under the ROC curve and balanced accuracy of separating remembered from forgotten trials using recall likelihoods aggregated across all saccades and electrode contacts for each trial. Both the area under the ROC Curve (0.801) and balanced accuracy (0.732) based on Shapley values-informed thresholds and directions exceeded all values in the control distribution generated by randomly sampling thresholds from the feature values. Taken together, Shapley values–informed thresholds and directions provided a valid and interpretable representation of SRND parameters underlying successful visual encoding.

To identify the brain region SREP characteristics most strongly influenced successful visual encoding, we computed the mean absolute Shapley value across trials for each electrode contact. This value quantifies the importance of each electrode contact's SRND parameters in predicting memory outcome. We ranked these values across anatomical regions to assess the spatial distribution of influence. **Figure 6 – Figure Supplement 1G** shows the mean absolute Shapley values for SRND parameters across cross-validation folds, which is consistent with feature importance based on Gini impurity reduction shown in **Figure 6G**. The concordance of feature importance across cross-validation folds exceeded all values from the null distribution (Kendall's coefficient of concordance = 0.966). As shown in **Figures 7G and H**, SREPs in the frontal and parietal regions had the strongest influence on visual encoding. To provide a finer spatial representation, we visualized the mean absolute Shapley values at the electrode contact level in **Figure 7 – Figure supplement 1C**. We projected the 3D parcellated brain regions and electrodes within these regions onto a 2D plane. Given the electrode size and approximately 1-cm reach of signals, we discretized this 2D space into a 5 mm × 5 mm grid. For each electrode,

we assigned its mean absolute Shapley value to the nearest grid element, and used a 1-cm spatial kernel to estimate the influence of SREPs on visual encoding for each grid element. This granular spatial map shows that SREP in frontal and parietal cortices and nearby white matter had the strongest influence on visual encoding, which closely mirrored the regional influence of SREP on visual encoding (**Figure 7H**).

Finally, we delineated the SREP profiles associated with successful visual encoding by mapping Shapley value–informed thresholds and directions for each SRND parameters (**Figure 7I–M**). SREPs with higher magnitude, especially in frontal and parietal regions, were associated with successful encoding (**Figure 7I**). This was further supported by the electrode contact-level SREP magnitude profile associated with successful visual encoding shown in **Figure 7 – Figure Supplement 1D**. Earlier SREP latencies across the cerebrum were generally associated with successful encoding (**Figure 7J, Figure 7 – Figure Supplement 1E**). Later peak times in most anatomical regions were associated with successful visual encoding (**Figure 7K**). At the electrode contact level, this association was consistent across electrode contacts in the left cortex, left white matter, and left limbic system, where later SREP peak times were associated with successful encoding (**Figure 7 – Figure Supplement 2A**). Oscillation power at 0.5 OCs before SREP latency showed hemispheric asymmetry: lower power in the left hemisphere and higher power in the right hemisphere were associated with successful encoding (**Figure 7L**). The electrode contact-level profile displays a region-dependent mapping between oscillation power and successful visual encoding (**Figure 7 – Figure Supplement 2B**). At 0.5 Ocs before SREP latency, oscillation phase near the trough-to-rising transition (e.g., $[-\pi, -\frac{\pi}{2}]$) was associated with successful encoding (**Figure 7M, Figure 7 – Figure Supplement 2C**). This pattern was observed for both polarities; however, because oscillation phase and polarity were dependent, the phase associated with successful encoding should be interpreted in the context of this dependence (**Figure 7 – Figure Supplement 2D–E, Figure 4D**).

These results delineate an SREP profile marked by early latency and strong amplitude, which was associated with successful visual encoding and likely reflects sensory readiness.

SREPs did not encode saccade direction

We have demonstrated that SREPs reflected neural states supporting visual encoding and characterized their profiles associated with successful visual encoding. An open question that follows is whether these SREPs align with the classical framework of saccadic corollary discharge, or whether they represent a novel saccade-associated mechanism. If SREPs were a narrowly

defined corollary discharge signal, they should convey information about saccade direction. We tested this using random forest regressors to examine whether SREP characteristics could predict saccade direction. Similar to the random forest model for classifying remembered vs. forgotten trials, the random forest regressor consists of an ensemble of decision trees, and each tree splits the data recursively based on feature thresholds. During training, the random forest regressor learned a set of features and split thresholds that minimize mean square error within each split node. After training, the output at each leaf node was the average target values across the training samples that reached that node. In our case, the target was the vector $[Sin(\theta), Cos(\theta)]$, where θ denotes the saccade direction in angular space, to account for the circular nature of saccade direction. This trigonometric transformation allows the use of Euclidean distance as a loss function and enables regression on a continuous 2D representation. We split the data at the saccade level, so that all SREPs recorded across electrode contacts during a saccade were assigned to either the training or the test set. This ensured that no information from the same saccade leaked into both sets. The input to the random forest regressor included the spatial location of electrodes and SRND parameters (e.g., SREP latency, peak-to-trough amplitude, peak time, phase, and power of preceding oscillatory activity). To encourage spatial generalization while preserving spatial specificity, we added Gaussian-distributed spatial jitter to electrode coordinates, which prevented the model from memorizing exact electrode contact locations. We obtained the predicted saccade direction using the two-argument arctangent function, $atan2$. **Figure 8A** illustrates the model architecture and training and evaluation pipeline.

To assess whether SREPs contained information about saccade direction, we compared the model's prediction error on held-out testing data to null distributions generated by permuting saccade directions. Across 5-fold CVs, the observed test mean squared error (MSE) for sin(θ) fell at the 64.6th percentile of the null distribution, and that for cos(θ) fell at the 93.0th percentile (**Figure 8B**). As a result, the mean absolute angular error between predicted and true saccade direction was higher than all values in the null distribution (**Figure 8C**). We next assessed if leveraging SREP characteristics across electrode contacts could reveal directional encoding. For each saccade, we obtained the prediction of saccade direction based on the predicted [sin(θ), cos(θ)] vectors averaged across electrodes (**Figure 8D**). We found that the testing mean absolute error (MAE) corresponded to the 52.2th percentile of the null distribution, indicating chance-level performance. To verify that this result was not due to suboptimal hyperparameter choices, we systematically evaluated the effect of key hyperparameters on model performance, including the number of trees, maximum tree depth, and minimum leaf size. Although increasing the maximum depth and decreasing the leaf size allowed the model to be more complex and better fit the training

data, the electrode contact-level and saccade-level model performance on the testing data remained consistently poor. This suggests that the failure to predict saccade direction was not due to model limitations but reflects the true absence of directional information in SREP features (**Figure 8 – Figure Supplement 1A**)

Subsequently, we examined whether SREPs in any contacts contained reliable directional information. We found that 5% of electrode contacts achieved a test MSE below the 5th percentile of the respective null distribution (**Figure 8E**). To identify the direction-informative electrode contacts, we compared the testing MSE with the 5th percentile of the respective null distribution, adjusted for the total number of electrode contacts. **Figure 8F** shows the ratio of saccade direction-informative electrode contacts across anatomical regions. No region exceeded a 3% ratio of direction-informative electrode contacts, except for the brainstem (13%, 1 of the eight micro wires near left hippocampus and brainstem showed directional information)**.** These results suggest that most, if not all, SRND parameters did not convey information about saccade direction.

To determine which SRND parameter showed directional modulation, we estimated the circular kernel density of the SRND parameters' values as a function of saccade direction (**Figure 8F**). For each SRND parameter, we quantified directional modulation as the non-uniformity of feature values across angular space, measured as the integral of the absolute deviation from the mean. **Figure 8G** shows the distribution of non-uniformity for the five SRND parameters. We compared them against the null distribution generated by shuffling SRND-direction pairings. For all five SRND parameters, the non-uniformity value in more than 95% of electrode contacts was below the 95th percentile of their respective null distributions (**Figure 8G**), which indicated a lack of systematic directional modulation.

We implemented a series of control models to completely rule out the presence of any systematic relationship between SREPs and saccade directions. First, we trained a fully connected, three-layer neural network to predict the sine and cosine of the saccade direction $Sin(\theta)$, $Cos(\theta)$, using SRND parameters and electrode coordinates (see **Star Methods**). **Figure 8 – Figure Supplement 1B** shows the network architecture. Each unit in the hidden layers was followed by a Rectified Linear Unit (ReLU) activation function to model non-linearity. The observed electrode contact-level mean squared errors (MSEs) for $Sin(\theta)$ and $Cos(\theta)$ were 0.924 and 1.035, corresponding to the 71.3rd and 27.9th percentiles of their respective null distributions (**Figure 8 – Figure Supplement 1C**). We obtained the null distributions of MSE from neural networks trained with permuted data. The MSEs for saccade directions were 10404.3 and 10262.9 deg^2 at the

electrode contact and saccade levels, corresponding to the 28.1st and 29.8th percentiles of the null distributions, respectively (**Figure 8 – Figure Supplement 1D**).

To ensure that the model training was not constrained by the trigonometric target representation, we implemented a separate neural network trained to minimize angular loss (**Figure 8 – Figure Supplement 1E**). This model yielded direction MSEs of 10662.2 and 10581.0 deg^2 at the electrode contact and saccade levels, corresponding to the 16.6th and 37.2nd percentiles of the null distributions, respectively (**Figure 8 – Figure Supplement 1F**). Finally, we implemented a von Mises kernel regression to model the saccade direction – SRND parameters relationship (see **Star Methods**). This model also failed to predict saccade direction above chance, with MSEs of 10782.5 and 10520.2 deg^2 at the electrode contact and saccade levels, corresponding to 22.0th and 22.2nd percentiles of the null distributions (**Figure 8 – Figure Supplement 1G**).

We next asked whether SRND parameters could inform a binary classification of saccade direction (left vs. right) using a random forest model. This reduces the complexity of the prediction problem and parallels our earlier prediction of successful visual encoding (**Figure 6C**). The binary random forest model performed at chance level. The electrode contact-level balanced accuracy was 0.501, corresponding to the 82.8th percentile of the null distribution. The saccade-level area under the ROC curve and balanced accuracy were 0.517 and 0.502, corresponding to the 51.6 and 73.0 percentiles of the null distribution, respectively. (**Figure 8 – Figure Supplement 1I**).

Across models and classification schemes, our results indicate the absence of a systematic relationship between SREPs and saccade directions. We next assessed whether SREPs reflected the motor planning for the subsequent saccade. Specifically, we trained and evaluated random forest models for predicting the direction of the next saccade based on SRND parameters. The electrode contact-level and saccade-level mean absolute errors were 89.863° and 89.613°. These errors corresponded to the 42.4 and 39.0 percentiles of the null distribution, respectively, suggesting that SREPs were unlikely to reflect prospective oculomotor planning.

These converging analyses demonstrate that SREPs do not convey directional information as a corollary discharge, but instead represent saccadic modulation supporting visual encoding.

Discussion

Here, we provided direct evidence that SREPs is an intrinsic generated neural mechanism in humans. We demonstrated an interplay between SREP waveforms and ongoing oscillatory activities, reinforcing the interpretation of these SREPs as saccadic modulation. Our explanatory analysis and machine learning models validated the hypothesis that SREPs have a role in visual encoding and confirmed the absence of directional tuning in SREPs. We identified SREP profiles associated with successful visual encodings. Active sensing mechanisms in other sensory systems or in other species coordinate sensing behavior with neural activities to optimize information processing[14,28,29]. Our findings demonstrate that an analogous mechanism exists to support human visual encoding. These findings position SREPs as potential targets for memory-enhancing neuromodulation and raise the prospect that incorporating biologically inspired modulation layers may lead to more intelligent artificial visual systems such as self-driving cars.

Presence, spatiotemporal structure, and behavioral relevance of SREPs

The novelty of our study is the shift from stimulus-locked responses to saccade-timed neural dynamics, showing that SREPs are present, spatiotemporally organized, and behaviorally relevant. Studies of neural mechanisms in human visual encoding have focused on brain responses to various visual stimuli[30,31]. Although SREPs were previously reported and shown to modulate neural activities in non-human primates, there is a dearth of direct evidence that these findings translate to humans[5,32,33]. This knowledge gap persists due to methodological constraints, such as the limited availability of human intracranial LFPs and the potential contamination of ocular-muscular artifacts in non-invasive electrophysiology. We leveraged a unique research setting in neurosurgical patients undergoing SEEG monitoring, and our findings provide direct evidence for the presence of SREPs in humans. These results corroborate and extend a recent study conducted by Staudigl et al., which showed that saccades coordinated amygdala-hippocampus interactions in humans[34]. Considered along with findings from non-human primates, the evidence suggests that saccadic modulation operates at multiple physiological scales, including neuronal spiking, evoked potentials, oscillation, and functional connectivity[2,29,35]. Likewise, we observed consistent SREPs, along with increased phase synchronization and increased oscillation power following saccades. Our findings that SREP waveforms are systematically associated with the phase of ongoing oscillations (**Figure 4**) suggest that neural dynamics of different scales at saccade times potentially reflect a latent state indexing attention or readiness for visual processing. Validating this hypothesis is important because, if correct, it would unify saccade-related neural phenomena into a coherent mechanistic framework. This

hypothesis makes two falsifiable predictions: (1) experimentally inducing saccades at different phases of ongoing oscillations affects other saccade-related neural dynamics, and (2) the encoding of other sensory information, such as auditory encoding, is enhanced when SREP profiles mirror those associated with successful visual encoding.

Previous studies of saccade-related neural dynamics have primarily focused on the visual system, where they were studied as mechanisms of visual perception, more specifically, perceptual stability[36,37]. More recent work has extended this scope, showing that saccades modulate neural activities beyond the visual system, including the medial temporal lobe[23], the thalamus[38], and the auditory cortex[16]. These findings suggest saccade-related neural dynamics may serve broader cognitive functions beyond visual perception. These studies primarily focused on individual brain regions, potentially missing coordinated saccadic modulation across the brain. Here, our study demonstrates widespread SREPs across the human brain, and reveals their systematic spatiotemporal organization through whole-brain spatiotemporal profiling. Earlier SREPs with greater amplitude emerged in the temporal lobe, a region implicated in high-order visual processing[39], followed by later SREPs across parietal and frontal regions (**Figure 5 – supplementary video**). This pattern suggests a cascading dynamic across regions that coordinate visual perceptual and encoding processes over time.

We interpret these results as follows. The early SREPs likely reflect that the temporal lobe actively prepares for the arrival of fovea inputs after each saccade. After these inputs are integrated into meaningful percepts, fronto-parietal networks mediate downstream encoding processes through information integration, working memory, and attention regulation[40,41]. Supporting this interpretation, SREPs in frontal and parietal regions exhibit the strongest influence on visual encoding performance (**Figure 7G and H**). One limitation of our dataset is the relatively low coverage of the early visual cortex, which constrains our ability to determine whether a latency gradient exists along the visual processing hierarchy. Even within this limitation, our findings provided convincing evidence that SREPs serve visual encoding, consistent with the view that human vision as an active sensing process[42]. Conceptually, our results invite a novel model of visual cognition. Instead of being a feedforward transformation of retinal input through the visual hierarchy, visual cognition involves neural events at saccade times that constitute a coordinated, whole-brain mechanism. We refer to these neural events as SREPs because, although our results show temporal consistency of LFP waveforms around saccades, it is correlational in nature. Without causal manipulation of saccades, it remains an open question whether saccades drive these neural dynamics or whether an underlying internal state coordinates both.

Implications for studying active visual encoding in humans

Our study has three primary implications for neuroscientific studies of active visual encoding in humans. First, we encourage research to adopt active-viewing paradigms and incorporate eye-tracking, as these approaches capture the dynamics of human visual processing in naturalistic settings. Second, SEEG recording in neurosurgical patients, such as those undergoing epilepsy monitoring, can reliably capture saccade-related neural dynamics that are uncontaminated by oculomuscular signals. Third, as saccades coincide with multiple processes, such as motor commands, shifts in retinal input, and cognitive processes, careful analytical dissociation is needed to attribute neural signals to behavioral and cognitive relevance. Our study showcases two key analytical strategies to address this challenge. We examined neural activities during saccades made on a blank screen to account for retinal input. This analysis reveals that these SREPs are largely non-visually mediated, consistent with previous animal studies[6,43]. Additionally, we separately correlated SREP characteristics recorded in intracranial EEG and EOG with saccadic behavior measures, such as saccade direction, thereby ruling out the contribution from motor signal. Below, we present additional control analyses and discuss their implications.

To quantify how long SREP waveforms were consistent across trials, we systematically varied the window size and compared the intertrial correlation distributions. We defined the pre-saccadic epoch as starting 1 second prior to saccade onset and with the same length as the peri-saccadic epoch (**Figure 2 – Figure Supplement 4A**). To quantify SREP consistency differences between pre- and peri-saccadic epochs, we computed the difference in the area under the cumulative distributions of the intertrial correlation curve (AUC) as follows: for correlations smaller than the crossing point, we calculated the difference by subtracting pre-saccade AUC from peri-saccade AUC; for correlations greater than the crossing point, we calculated the difference by subtracting peri-saccade AUC from pre-saccade AUC. Aligning with the main findings, we found that SREPs were more consistent than pre-saccadic LFPs (**Figure 2 – Figure Supplement 4B**, $t = 5.349$, $p < 0.001$, Cohen's $d = 0.274$, $n = 382$, two-sided t-test). This result is consistent across window length (**Figure 2 – Figure Supplement 4B)**. Interestingly, distributions of intertrial correlations were less normally distributed when the analysis windows were shorter, for both pre- and peri-saccadic EP (**Figure 2 – Figure Supplement 4B**).

These results have several implications. First, SREPs are transient. As we increased the window length, we incorporated ongoing activity that is not related to the saccade, therefore reduced the waveform consistency. The same pattern observed in the pre-saccade neural field potential may suggest temporal coordination between visual cognitive processes and motor

rhythms (i.e., saccades). Our observed oscillation phase clustering preceding saccades supports this view (**Figure 4D**). Alternatively, the window length influences intertrial correlation distribution. Given that neural field potential is structured in time, with features like long-range correlations, shorter windows are more likely to yield extreme intertrial correlation values[44]. To further validate our results, we performed an additional surrogate analysis by shuffling trial data to create artificial pre- and peri-saccadic LFP waveforms. We found that the separation effect size, which measures how peri-saccadic LFP differs from baseline LFP, was significantly reduced in surrogate data compared to real data ($t = -65.086$, $p < 0.001$, Cohen's $d = -1.804$, two-sided t-test, $n = 2604$, **Figure 2 – Figure Supplement 4C**). The above analyses suggest that the LFP waveform following saccades is not spontaneous but represents consistent saccadic modulation. This conclusion is further supported by our finding that SREPs predicted successful visual encoding.

We evaluated the generalizability of the model across individual subjects by assessing electrode contact- and saccade-level balanced accuracy averaged across cross-validation iterations, separately for each subject's test data. For each subject, we generated a null distribution of balanced accuracy using models trained on permuted labels. **Figure 6 – Figure Supplement 1E** shows the percentile rank of balanced accuracy relative to the null distribution for each subject. The observed saccade-level balanced accuracy exceeded the null distribution for 6 out of 8 subjects, demonstrating relatively good generalizability. For the remaining two subjects, saccade-level balanced accuracy corresponded to the 83rd and 72nd percentiles of the null distribution, indicating performance above chance but with limited generalizability. This is not unexpected. Although the full cohort provided broad coverage across the cerebrum, electrode coverage for individual patients was often more limited, for example, restricted to a single hemisphere depending on the suspected seizure onset zone. We also note that these two subjects had the fewest valid remembered and tested visual encoding trials (12 and 14), which might limit the model's ability to fully learn the mapping between SRND parameters and successful encoding at the brain regions sampled in these individuals.

Because the anatomical organization of brain regions and their associated functions is highly complex and not expected to follow a linear mapping, we used non-linear models (i.e., random forest models), to learn the mapping from SRND parameters recorded at specific brain regions to visual encoding. We were interested in whether there is a linear relationship between the SRND parameters and visual encoding. For this purpose, we trained logistic regression models with the same input and labels as in training random forest models, except we used the absolute peak-to-trough amplitude, as the polarity was determined by the preceding oscillation

phase. The logistic regression performed well above chance, but not as well as the random forest model (**Figure 6 – Figure Supplement 1I**). For testing data, the electrode contact-level balanced accuracy was 53.1%, the saccade-level ROC AUC was 63.2%, and the saccade-level balanced accuracy was 59.5%. Under the assumption of a linear relationship, the weight coefficient for each SRND parameters shows that a higher absolute SREP peak-to-trough value was associated with remembered trials, which aligns with the mapping identified in the random forest models (**Figure 7I**).

Proposed mechanisms of SREP and their relationship with successful visual encoding

Our investigation was motivated by Katz et al., who reported saccade-related inhibition of single units in mnemonic structures and interpreted it as a corollary discharge-like signal[2]. While our study built upon their findings, our results reveal a distinct form of saccadic modulation. Specifically, our study provides direct evidence for the role of SREPs in visual encoding. Moreover, the observed SREPs did not systematically encode saccade direction, which is a hallmark of corollary discharge. Our findings thus suggest that neural dynamics following saccades are another form of saccadic modulation that is more consistent with the active sensing framework and the study by Leszczynski et al. in non-human primates[16]. The existing literature and our results suggest two mechanistically distinct forms of saccadic modulation, including an early corollary discharge and a later modulation supporting visual encoding[45–47].

Within the active sensing frameworks, a theory proposed that saccades modulate sensory processing by resetting ongoing oscillations to a high-excitability phase[3]. Our results suggest an alternative, but not mutually exclusive, mechanism of SREPs. We observed SREPs with positive and negative polarities at the same recording sites, and their waveform depends on the ongoing oscillation phase. As these SREPs were derived from locally re-referenced field potential, they reflected the mesoscale current flows driven by synchronized synaptic currents[48]. Based on this, we propose a model in which each brain region follows its intrinsic rhythm, and at saccade times, neural activities are synchronized in a way that is contingent on the local rhythm. The consistent LFP waveforms time-locked to saccades may result from the temporal coordination between saccade and intrinsic brain rhythms. This interpretation is supported by observations of phase clustering before saccades shown in this study (**Figure 4**) and in other studies[23,49]. This model predicts higher intertrial coherence and elevated oscillatory power at saccade times, which are also validated in our study and in previous works[3,16]. Given that SREP is a brain-wide phenomenon, a global, uniform phase reset would likely interfere with ongoing computations, for

example, information coding[50]. Thus, a phase-contingent synchronization mechanism for processing incoming input may be functionally more beneficial.

All visual inputs are not remembered equally[51]. Our study shows that SREP characteristics and ongoing oscillations explain this variability. We distilled the SRND parameter–encoding relationship learned by the random forest model into a lower-dimensional space consisting of thresholds and direction rules based on Shapley values. The extrapolated mapping predicted successful visual encoding slightly better than the random forest classifier, suggesting that this representation preserves and identifies a more generalized mapping between visual encoding and SREPs. The lower-dimensional representation lay the foundation for interpreting the relationship between SRND parameters and visual encoding. Specifically, successful encoding was associated with greater SREPs amplitude, particularly in frontal and parietal regions (**Figure 7I**). It is our interpretation that stronger SREPs represent greater attention and may signal priority for encoding in memory circuits[52]. In addition, successful encoding is associated with earlier SREPs in general, which mirrors Staudigl et al.'s finding that neuronal responses in medial temporal lobe associated with remembered images occurred sooner than later forgotten images[34]. These findings suggest that earlier saccadic modulation is beneficial for visual encoding, although the exact mechanism remains to be elucidated.

Previous studies have demonstrated that greater oscillation power in the hippocampus was associated with better recognition memory in non-human primates[3]. Our results provide a whole-brain mapping between oscillatory activities and successful visual encoding. The normalized oscillation power at 0.5 OCs before SREP latency showed hemispheric asymmetry. In general, lower preceding power in the left hemisphere and higher power in the right were associated with successful encoding. If SREPs represent synchronization, contingent on local rhythm, then regions where higher preceding oscillation power predicts successful encoding likely belong to the encoding-relevant network. In other words, their ongoing rhythmic activity is functionally aligned with visual encoding. We caution that the functional meaning of oscillation is frequency- and region-dependent. For example, increased posterior alpha in visual cortex has been linked to enhanced signal-to-noise ratio via selective inhibition of background activity[53], whereas increased hippocampal theta is associated with active information encoding[54]. Given that the dominant frequency across cerebrum lies in the theta–alpha range (5–12 Hz; **Figure 4 – Figure Supplement 1D**), it is plausible that both inhibitory gating and encoding mechanisms are at play.

Resource availability

Lead contact

Requests for further information and resources should be directed to, and will be fulfilled by, the lead contact, Peter Brunner (pbrunner@wustl.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All analysis scripts and codes for plotting individual figure panels are available on a GitHub repository (https://github.com/GanshengT). Processed data used in the analyses are archived on Zenodo data repository. Raw imaging and electrophysiological data can be de-identified and provided to qualified researchers upon request due to the large file size (~95 TB) and in accordance with our IRB-approved consent. All processing codes, example anonymized raw data, and a script illustrating the data-processing pipeline are available on a GitHub repository (https://github.com/GanshengT).).

Acknowledgements

This work was supported by NIH/NIMH (R01-MH120194), NIH/NIBIB (P41-EB018783, R01-EB026439), McDonnell Center for Systems Neuroscience, Fondazione Neurone, NIH/NINDS (U24-NS109103, U01-NS108916, U01-NS128612, R21-NS128307), and NIH/NEI (R01-EY037195). The manuscript was edited by the Scientific Editing Service of the Institute of Clinical and Translational Sciences at Washington University, which is supported by an NIH Clinical and Translational Science Award (UL1 TR002345).

Author contributions

Conceptualization: G.T.; Data curation: G.T., P.B., and J.T.W.; Formal analysis: G.T.; Funding acquisition: E.C,L., P.B., J.T.W., C.S.I., J.R.M., and S.B.H.; Investigation: G.T., P.B., J.T.W., and E.C.L.; Methodology: G.T., P.D., H.C., Y.L., and Z.L.; Software: G.T., P.B., and J.R.S.; Supervision: P.B., J.T.W., and E.C.L.; Validation: G.T., P.B., H.C., P.E.C., S.S.S., H.P., X.L., Z.L., and E.C.L.; Visualization: G.T., P.D., and P.B.; Writing – original draft: G.T.; Writing – review & editing: G.T., J.T.W., P.B., and E.C., Task design: J.R.S., K.L.W., M.K.H., and J.M.C.

Declaration of interest

The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process.

During the preparation of this work, the author(s) used ChatGPT and Grammarly AI in order to improve the readability of the manuscript. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

STAR Methods

Experimental model and study participant details

Eight epilepsy patients (two males and six females) who were undergoing intracranial monitoring with stereoelectroencephalography (SEEG) at Washington University School of Medicine participated in this study. All participants provided written informed consent in accordance with protocols approved by the Institutional Review Board (IRB protocol number: 202104033-1001). Patients were eligible if they were diagnosed with epilepsy and scheduled to undergo intracranial video monitoring for seizure onset localization. None of the include patients reported subjective cognitive impairment, and no abnormalities in eye movements were noted during clinical evaluation. Patient demographics and seizure diagnosis are described in **Supplementary Table 2 Patient Characteristics**.

Electrode implantation and clinical procedure

All patients underwent implantation of multi-contact depth electrodes (sEEG Depthalon depth electrode, PMT® Corporation, MN, USA) for seizure monitoring. Each electrode shaft had 8–16 contacts (2 mm cylindrical contacts length, 0.8 mm diameter, spaced 3.5 mm apart center-to-center). In 6 patients, two hybrid depth electrodes with macro- and micro-contacts (Behnke Fried Depth Electrode, Ad-Tech® Medical Instrument Corporation) were implanted. Each hybrid depth electrode has a microwire bundle in addition to 8 standard macro contacts. The microwire bundle consists of nine 40-μm diameter wires, which were trimmed such that they extend up to 5 mm beyond the tip of the macroelectrode during the implantation. Eight of the microwires are coated in polyamide insulation up to the tip, resulting in an impedance of 100–300 kΩ; the ninth wire is uninsulated, serving as a reference.

The clinical team planned the electrode placement to ensure appropriate coverage of the brain for seizure localization. T1-weighted imaging (MPRAGE sequence, voxel size $1.0 \times 1.0 \times 1.0\ mm^3$) was obtained in the presurgical workup and loaded into the planning workstation. The neurosurgeon implanted electrodes, guided by a stereotactic robotic system (ROSA ONE Brain, ZIMMER BIOMET, IN, USA). After the surgery, patients underwent a CT scan,

which was co-registered to their pre-operative MRI to obtain precise 3D coordinates of every electrode contact. We describe the electrode localization method in detail in the **Method Details** section below. Patients were then awakened from general anesthesia and transferred to the Epilepsy Monitoring Unit (EMU) for long-term intracranial electroencephalography (iEEG) monitoring. The duration of EMU stay varied by patient, typically lasting several days to weeks, depending on seizure frequency. During this period, the research team introduced the study to the patients and obtained written informed consent before beginning any experimental recordings. After consent, research sessions were scheduled in coordination with the clinical team to ensure there were no disruptions to standard clinical care. Standard clinical protocols may taper or withdraw antiseizure medications to provoke seizures during iEEG monitoring. In some cases, partial sleep deprivation was employed to increase seizure likelihood.

Method details

Study design

This study is part of a broader protocol (registered clinical trial identifier: NCT05065450) investigating how amygdala stimulation affects memory formation. To assess recognition memory, subjects were asked to perform visual encoding tasks in which they viewed a series of images of everyday objects or scenes from a standard image set[55]. Memory for these images was assessed approximately 24 hours later. Depending on the patient's availability and length of stay in the EMU, they could elect to complete multiple iterations of the study protocol across different days. Each study iteration followed a two-day structure, comprising a visual encoding task on Day 1 and a recognition memory test on Day 2. The encoding task on Day 1 was administered into two or four sessions per participant to mitigate fatigue and accommodate clinical scheduling. The encoding task comprises 180 trials in total, where a trial is defined as the presentation of a single image. The task structure remained constant across iterations. Each study instance used a unique set of images. The details of the behavioral paradigm are provided below.

In less than half of the trials, one second of electrical stimulation was delivered to the amygdala following presentation of each image. The current study focuses on normative brain activity around saccade times during visual encoding. To this end, we only included trials that did not involve any brain stimulation. Although object and scene processing might engage dissociable neural mechanisms, our primary aim was to characterize neural dynamics time-locked to saccades; therefore, we pooled object and scene trials. Categories were balanced across the set, and exploratory checks did not indicate category-dependent differences in the saccade metrics. We defined valid visual encoding trials for subsequent neural analyses based on two criteria: 1.

Subjects maintained their gaze on the monitor for more than half the duration of this trial, based on the eye-tracking data; 2. No brain stimulation was delivered during this trial. These decisions yield a total of 1130 valid visual encoding trials and 7292 valid saccades across eight patients. The detailed description of the data set, such as the total number of study iterations per participant and the total number of encoding trials, is summarized in **Supplementary Table 1**.

Behavioral paradigm

Participants completed visual encoding tasks at their bedside during their stay in the EMU. Before the main task began, participants received standardized instructions and completed a brief training session to become familiar with the task structure. The images presented during training were not repeated during the main task, and data from the training session were not included in the analysis.

After training, participants began the encoding task. A series of trial-unique color images was presented on a 27-inch monitor (TUF GAMING VG27AQL1A, ASUS, USA) mounted on a mobile cart positioned next to the patient's bed. Each image was displayed at the center of the screen for 3 seconds, followed by a 6-second inter-stimulus interval (ISI). The image occupied 5.56 × 5.56 inches on the screen at a viewing distance of approximately 50 cm. Participants were instructed to freely view each image and remember it for a later memory test. A fixation cross was presented during the ISI, but participants were not required to fixate on it. In brain stimulation trials, 1 second of electrical stimulation was delivered to the amygdala during the first second of the 6-second ISI. Only non-stimulation trials were included in the current analysis. Approximately 24 hours after completing the encoding task, participants performed a self-paced recognition memory test. The test consisted of 89% previously seen images and 80 novel images, randomly intermixed. Each image was displayed for a mandatory 1-second viewing period, after which participants used a keyboard to indicate whether they had seen the exact image before ("old" vs. "new"). No feedback was given during the duration of the memory test.

Experimental setup and data acquisition

We used a desktop computer (ASUS ROG STRIX B550-E GAMING motherboard; AMD Ryzen 9 5900X 12-core processor; 32 GB DDR4 RAM; NVIDIA GeForce RTX 3070 GPU) running the BCI2000 platform to manage stimulus presentation, behavioral logging (e.g., keyboard responses), and data synchronization[20]. We recorded and synchronized photodiode signals, eye-tracking data, and electrophysiological data. An 27-inch monitor (TUF GAMING VG27AQL1A, ASUS) with a resolution of 2560 × 1440 pixels and a 120 Hz refresh rate was connected to the desktop computer and mounted to a mobile cart using an adjustable monitor arm, allowing flexible

positioning to accommodate each patient's posture and bed orientation. The monitor was placed at approximately eye level, about 50 cm from the participant's eyes.

A screen-based eye tracker (Tobii Pro Fusion, Tobii AB, Sweden) was mounted at the bottom edge of the monitor and recorded gaze position at 60 Hz. We performed eye-tracker calibration at the start of each task using *Tobii Pro Eye Tracker Manager* software, and repeated as needed if the relative position between the participant's eyes and the eye tracker changed.

We recorded iEEG data using a clinical SEEG system (Neurofax EEG-1200, Nihon Kohden, Irvine, CA) with a JE-120 amplifier (256 channels). Signals from depth electrode contacts were sampled at 2,000 Hz with 16-bit resolution. Per standard clinical practice, two subgaleal electrodes served as the reference and ground during signal acquisition. Simultaneous scalp EEG was recorded as part of the clinical montage. We specifically used frontopolar electrodes (Fp1, Fp2) to detect oculomotor artifacts.

To capture accurate timing of visual stimulus onset, a photodiode (g.tec medical engineering GmbH, Austria) was affixed to the bottom-left corner of the screen. At the onset of image presentation, a white dot was shown at the bottom-left corner of the screen, aligned with the photodiode. The white dot disappeared at the onset of the ISI. The photodiode output was routed through a multimodal trigger conditioner box (g.TRIGbox, g.tec medical engineering GmbH, Austria), and synchronized with electrophysiological and eye-tracking recordings.

Anatomical localization of the electrode

We localized electrodes based on post-operative CT scans co-registered to pre-operative T1-weighted MRI scans (0.5 mm isotropic resolution). To co-register the CT volume to the MRI volume, we first converted the MRI and CT scans from DICOM to NIfTI format using *dicm2ni*, a tool from NeuroImaging Tools and Resources Collaboratory. We rigidly transformed the CT volume to place its geometric center at the world coordinate origin. This step facilitated robust rigid-body coregistration using the *SPM12* mutual-information-based registration. For each subject, the pre-operative MRI was treated as the anatomical reference, and the CT image was coregistered to it.

We first manually localize macroelectrode contacts using Freeview, which is the image viewing program of the FreeSurfer software package. To improve precision and reproducibility, we implemented a semi-automated algorithm to align the manually identified contact positions with the most intense local signal in the CT volume, indicative of metal artifact. For each manually identified macro contact, we extracted a local CT subvolume (5*5*5 voxels, corresponding to

approximately 16 mm^3) and performed a data-driven intensity thresholding procedure. Starting from an empirical threshold, estimated via kernel density estimation of voxel intensities, the algorithm progressively lowered the threshold and identified voxels with signal exceeding this threshold, until a connected component exceeding the expected physical volume of the electrode was found. Based on manufacturer specifications, the expected physical volume was a cylinder with a 1.28-mm diameter and a 1.57-mm length. We assigned the centroid of this component as the refined electrode location and validated the refined locations using Freeview. We then inferred microelectrode positions based on the geometry of the macro contacts. We fitted a line for each lead based on macroelectrode positions via singular value decomposition. The initial estimated micro contact was 3 mm beyond the most proximal macro contact, along the same shank axis. The value, 3 mm, was informed by the surgeon who cut the micro wires during planning. We localized the most intense cluster in the CT signal around the initial estimation microwire position, using the same intensity-guided connected-component algorithm.

To further verify the accuracy of electrode localization, we exported the executed electrode trajectories from the Robotic Operating Surgical Assistant (ROSA) platform, where trajectories were defined in the coordinate space of the intraoperative imaging system. To transfer these trajectories into the same anatomical reference space as the localized electrode contacts, we performed rigid-body image coregistration between the ROSA display image and the pre-operative T1-weighted MRI. The transformation was then applied to the extracted ROSA trajectories. As a result, both the planned lead trajectories and the electrode localizations were in the same right-anterior-superior (RAS) coordinate space defined by the pre-operative MRI.

To identify the brain regions associated with each electrode, we performed brain segmentation using FreeSurfer (v7.4.1), which parcellates cortical and subcortical structures based on the Desikan-Killiany atlas. We assigned anatomical labels to each electrode contact using three approaches. First, we applied a majority-vote volumetric method, in which each electrode was assigned to the brain structure encompassing the majority of segmented voxels within a 10 mm Euclidean radius from the contact center. This accounts for the spatial reach of macroelectrodes. Second, we used a surface-based inclusion method, where electrodes were labeled according to whether their coordinates were enclosed within the triangulated surface mesh of a segmented structure. Third, we implemented a probabilistic labeling approach, in which each structure was assigned a probability proportional to the number of its voxels within a 10 mm sphere centered on the electrode. Without further specification, the anatomical label in this study derived from the first method.

For group-level analysis and visualization, we mapped individual electrode locations to a common anatomical space defined by the native MRI of a reference subject. To achieve this, we used the ANTs (Advanced Normalization Tools) image registration framework to coregister each subject's preoperative T1-weighted MRI to the reference MRI. We applied only the affine component of the transformation to the electrode coordinates to preserve their linear relationships. The affine matrices were estimated within the ANTs' non-linear optimization pipeline to account for anatomical variability.

To quantify sampling density across brain regions, we calculated the weighted count of electrode contacts per brain region (**Figure 1D**). In other words, when electrode contacts (~2 mm^3 volume) overlapped multiple brain regions, their contributions were proportionally allocated among these regions. We further quantified sampling density per region by dividing the weighted contact count within each brain region by the volume of that region (**Figure 1 – Figure Supplement 2**).

Eye-to-Electrode Distance calculation

To assess if the observed SREP could be attributed to oculomotor signals, we performed regression analysis between the eye-to-electrode distance and SREP characteristics such as peak amplitude. We identified the center of each eyeball for each subject based on their pre-operative MRI scans and then computed the 3-dimensional (3D) Euclidean distance from the eyes to each electrode contact's coordinates. We used the minimum of these two distances, representing the distance from an electrode to the closest eye, as the eye-to-electrode distance for subsequent analyses.

Intracranial EEG preprocessing

To ensure the analysis focused on physiologically meaningful, non-pathological signals, we systematically identified and excluded electrodes exhibiting epileptiform discharges or artifactual activity (e.g., due to poor contact). To identify electrodes with artifactual activity, we computed the root-mean-square (RMS) amplitude of the 60 Hz line-noise component using a narrowband infinite impulse response (IIR) filter using *iirpeak* in MATLAB with a peak frequency of 60 Hz and a bandwidth of 0.1 Hz as inputs. Electrode contacts with 60 Hz RMS amplitude deviating more than two standard deviations (SDs) from the median were flagged and excluded. To detect potential shorted (i.e., electrically bridged) contacts, we computed pairwise Pearson correlation coefficients between all contacts within each electrode trajectory after re-referencing. For re-referencing, we applied common average referencing for each electrode trajectory after excluding

electrode contacts with excessive 60 Hz line noise. We identified contacts exhibiting excessively high correlation with non-adjacent neighbors ($R^2 > 0.999$) as potentially shorted and excluded them from further analysis. We did not find any shorted contacts.

To identify electrodes with epileptiform activity, we first extracted high-frequency iEEG signals (>250 Hz) using a 4th-order Butterworth filter (*butter* function in MATLAB). We computed the analytic envelope of the high-pass filtered signal using the Hilbert transform (*hilbert* function in MATLAB) and identified time points where the envelope exceeded five SDs above the mean. We identified time points where both the raw signal amplitude and its temporal gradient exceeded 5 SDs above their means. We marked the identified time points as epileptic and calculated the proportion of epileptic time per electrode. We excluded electrodes exhibiting epileptiform activity for more than 1% of the recording duration. We applied the procedures for identifying line noise and shunted contacts to microwire recordings. We did not apply the epileptiform activity exclusion criterion to microwires, as they recorded both neuronal spiking activity and LFPs. **Supplementary Table 3** summarizes per-subject counts of retained electrodes for subsequent signal processing.

To enhance spatial specificity and remove common artifactual noises, we re-referenced intracranial signals using a small Laplacian re-referencing, separately for macro electrode and micro wires. We calculated inter-contact Euclidean distances derived from the three-dimensional electrode coordinates. For each contact, neighboring electrodes were identified as those located within a predefined spatial threshold of 6 mm. We then epoched the resulting Laplacian re-referenced signals relative to stimulus onset for subsequent SREP analyses. To ensure the robustness of downstream analyses to the referencing scheme, we kept track of the number of neighboring electrodes. For some electrodes located at the ends of a shank, where only a single adjacent contact was available, the small Laplacian re-referencing reduced to bipolar referencing.

Saccade detection and characterization

The recorded eye tracking data included three-dimensional eye position and two-dimensional gaze direction, expressed relative to the top-left corner of the display monitor. The Z-axis was orthogonal to the screen plane, so that all gaze positions had $Z = 0$, and the eye position's Z-coordinate represented its distance from the screen. We preprocessed the eye-tracking data separately for each eye. We marked eye tracking samples as invalid if they met two criteria: (1) the eye tracker reported loss of tracking (e.g., during blinks or partial eyelid closures), or (2) the gaze coordinates were outside the physical screen bounds. To account for signal artifact at the edges of blinks or tracking dropouts, we extended each invalid segment by ±30 ms on both sides.

We computed the proportion of valid samples for each eye, and selected the eye with the higher proportion of valid data for further analysis. We time-locked the gaze data to the image presentation. We defined each trial as the interval from 2 s before the onset of the fixation cross presentation to 2 s after image presentation offset (**Figure 1b**). To ensure that our analyses reflected periods when subjects were encoding visual information, we excluded individual trials with more than 50% invalid eye tracking data.

We segmented the continuous gaze traces into periods of fixation and saccades using the *ClusterFix* algorithm[56]. In brief, the algorithm identified eye movement events by clustering consecutive gaze coordinate samples based on distance, velocity, acceleration, and angular velocity to classify them as either fixations or saccades. We linearly interpolated invalid samples in gaze data to maintain the temporal continuity required by the clustering algorithm. After clustering, we verified that both onset and offset for each identified saccade and fixation occurred during valid (e.g., noninterpolated) samples. Saccadic or fixation events overlapping interpolated segments were flagged as invalid and excluded from downstream analyses. This conservative approach ensured that only fixations and saccades derived from true gaze samples were included in the analysis. As a validation, we compared the velocity and duration of detected saccades and fixation with expected human eye movement parameters (**Figure 1 – Figure Supplement 1**).

We calculated saccade eccentricity, saccade direction, and saccade and fixation duration based on the onset and offset time of the detected saccade or fixation. We transformed gaze coordinates into a vector representing gaze coordinates relative to eye location in the right-anterior-superior (RAS) coordinate system, and normalized it to unit length. Based on the normalized vector $\overrightarrow{(gaze_x, gaze_y, gaze_z)}$, we calculated the azimuth and elevation angle ($\theta_{azimuth}$ and $\theta_{elevation}$) in polar coordinates.

$$\theta_{azimuth} = tan^{-1}(\frac{gaze_y}{gaze_x})$$

$$\theta_{elevation} = tan^{-1}(\frac{gaze_z}{\sqrt{gaze_x^2 + gaze_y^2}})$$

we calculated the angular change $\Delta\theta_{azimuth}$ and $\Delta\theta_{elevation}$ based on $\theta_{azimuth}$ and $\theta_{elevation}$ at saccade onsets and offsets. We defined the saccade eccentricity as $\sqrt{\Delta\theta_{azimuth} + \Delta\theta_{elevation}}$. We defined saccade direction θ_{dir} as the orientation of the 2D angular displacement vector connecting gaze positions before and after each saccade. In this convention,

0 rad (e.g., 0°) corresponds to rightward saccades, $\frac{\pi}{2}\,rad$ (e.g., 90°) to upward saccades, $\pi\,rad$ (e.g., 180°) to leftward saccades, and $\frac{3\pi}{2}\,rad$ (e.g., 270°) to downward saccades.

$$\theta_{dir} = tan^{-1}(\frac{\Delta\theta_{elevation}}{\Delta\theta_{azimuth}})$$

To study whether SREPs encode the direction of the next saccades, we identified consecutive saccades that occurred with both spatial and temporal continuity constraints. Specifically, we required the spatial distance between the end of the current saccade and the start of the subsequent saccade to be less than 20% of the display dimension, and the inter-saccadic interval to be shorter than 550 ms.

Identification of saccade-related evoked potentials (SREPs)

To identify SREPs, we analyzed re-referenced LFP signals time-locked to saccades. The dataset was organized hierarchically: each subject completed one or more encoding tasks, each task consisted of multiple recording sessions, and each session contained LFP signals from multiple electrode contacts. Within each session, there were multiple trials (e.g., image presentation), during which one or more saccades occurred. For each encoding task completed by a subject, we segmented the re-referenced signals into epochs relative to saccade onsets and combined all epochs across trials and sessions within that task. This study focused on saccades occurring during the encoding of the presented image. To isolate SREPs from stimulus-onset responses, we excluded saccades occurring 200 ms before or after the image onset. For control analyses, we separately analyzed saccades made during fixation-cross presentation, when the screen background was uniformly black with a central white fixation cross occupying approximately 2% of the screen's horizontal dimension. The fixation-cross presentation period provided a baseline condition for evaluating whether peri-saccadic EPs reflected visual transients. For each epoch, we defined the pre-saccadic interval from -200 to -10 ms relative to the saccade onset as the baseline period. We normalized each epoch by z-scoring its LFP signal based on the baseline mean and SD. To obtain the mean SREP waveform for each electrode contact, we averaged the normalized LFP signals across all saccades, trials, and sessions.

To determine whether SREPs were consistent across epochs, we calculated the pairwise Pearson correlation coefficients among normalized epochs within the peri-saccadic window (-10 to +40 ms relative to saccade onset) and, for comparison, within the baseline window[22]. These

windows were defined to capture the immediate neural dynamic following a saccade while avoiding temporal overlap between EPs arising from consecutive saccades. To ensure robustness of subsequent analyses, we systematically varied the baseline window length and confirmed consistent results (**Figure 2—Figure Supplement 4**). Unless otherwise noted, the results presented in this study were derived using the −200 to −10 ms baseline.

The distribution of correlation coefficients during the peri-saccadic period exhibited a bimodal pattern, suggesting SREPs of two polarities (**Figure 2d**). To statistically evaluate this observation, we quantified how much the peri-saccadic correlation distribution deviated from the unimodal baseline distribution. We quantified bimodality based on the separation between the inter-trial correlation distributions during the peri-saccadic and baseline periods. We computed the intersection point of their cumulative distributions and measured absolute difference in area under the curve (AUC) between them as their separation effect size (**Figure 2 – Figure Supplement 2**). This separation effect size reflected the extent to which inter-trial correlations during the peri-saccadic period were bimodal, with larger values indicating stronger bimodality. To assess statistical significance, we applied Kolmogorov–Smirnov tests to determine the separation effect size threshold corresponding to the significance level corrected for multiple comparisons.

This analysis confirmed the presence of bimodal SREPs. To identify the two types of SREPs, we applied hierarchical clustering to the inter-trial correlation matrix using the MATLAB *linkage* and *cluster* functions with *maxclust* input set to 2. Each element of this matrix represented the correlation coefficient between a pair of epochs. In brief, hierarchical clustering groups epochs based on their pairwise correlation, and then iteratively merges those with the most similar waveforms to form a dendrogram that captures the hierarchical relationship among epochs. To ensure the robustness of the hierarchical clustering algorithm, we required a minimum cluster size of 25% of the total number of epochs. Failure to meet this criterion could occur when the inter-trial correlation structure was weakly bimodal. In this study, only three electrode contacts had peri-saccadic correlation matrices that failed to meet this criterion. For these electrode contacts, we implemented an adaptive re-clustering procedure that iteratively removed epochs with globally weak correlations, defined as the lowest mean absolute correlation across epochs, and repeated the clustering until both clusters satisfied the size criterion.

After identifying SREPs of two polarities, we quantified the separability between peri-saccadic and baseline correlation distributions for each cluster. To do so, we computed the receiver operating characteristic (ROC) curve, which characterizes how well the two distributions can be distinguished. The ROC curve plots the true positive rate (TPR), defined as the fraction of

peri-saccadic correlations exceeding a given threshold, against the false positive rate (FPR), defined as the fraction of baseline correlations exceeding the same threshold. Thresholds were sampled uniformly between −1 and 1. The area under the ROC curve (AUC) measures the separability. Larger AUC values indicate greater divergence of peri-saccadic correlation structure from baseline, reflecting more consistent SREP waveforms.

As a control, we repeated the same clustering and separability quantification procedures after permutating peri-saccadic and baseline correlation matrices. We applied the identical adaptive hierarchical clustering procedure on the baseline correlation matrix. We then recomputed the ROC curve following the same procedure, but with the roles of peri-saccadic and baseline correlations reversed. In other words, the TPR now represented the fraction of baseline correlations exceeding a given threshold, and the FPR represented the fraction of peri-saccadic correlations exceeding the same threshold. We then compared the observed and control AUC values to identify electrode contacts with consistent SREPs, using AUC difference as a measure of consistency. This control accounts for potential bias induced by the clustering algorithm and allows for the estimation of separability specifically attributed to peri-saccadic LFP waveform similarity.

Quantification of SREP consistency and regional response ratio

To identify the brain region where LFP waveforms in the peri-saccadic window were most consistent, we computed the response ratio for each brain structure. The response ratio was defined as the proportion of electrode contacts whose peri-saccadic AUC exceeded the mean + 2 SDs of their respective baseline AUC values (**Figure 2G**). The use of proportion accounts for nonuniform sampling density across the cerebrum. We calculated the response ratio for the following structures: White Matter (Left, Right), Hippocampus (Left, Right), Amygdala (Left, Right), Basal Ganglia, Thalamus, Orbitofrontal (Left, Right), Inferior Frontal (Left, Right), Middle Frontal (Left, Right), Superior Frontal (Left, Right), Anterior Cingulate, Posterior Cingulate, Motor (Left, Right), Anterior Temporal (Left, Right), Middle Temporal (Left, Right), Superior Temporal (Left, Right), Inferior Temporal (Left, Right), Medial Temporal (Left, Right), Fusiform (Left, Right), Parietal (Left, Right), Occipital (Left, Right), and Insula (Left, Right).

To robustly quantify the fraction of electrode contacts within a region that exhibit consistent saccade-related activity, we systematically varied the responsiveness threshold from mean + 1 SD to mean + 4 SD of baseline AUC (**Figure 2—Figure Supplement 1**). For each threshold, we calculated the proportion of electrode contacts exceeding the threshold. This yielded a curve where the y-axis represented the response ratio defined before, and the x-axis represented the

responsiveness threshold. As an integrated measure of regional LFP waveform consistency following saccades, we computed the area under this curve (AUC) across all thresholds. As a control for the number of sampling electrodes in different brain structures, we resampled electrodes with replacement to equalize sample sizes across regions, then repeated the calculation of the AUC. We computed the Pearson correlation coefficient between AUC with and without resampling to confirm trivial effects on sampling density.

To determine the spatial distribution of the degree of LFP waveforms consistency following saccades, we computed the difference in AUC between peri-saccadic and baseline inter-trial correlation distributions (**Figure 2–Figure Supplement 2c**). As a complementary measure, we defined the degree of consistency as the average AUC of the probability density over inter-trial correlation following saccades, computed separately for the two clusters and then averaged (**Figure 2–Figure Supplement 2D**).

Characterization of SREP waveforms

To test whether the observed SREPs represented oculomuscular signals or neural mechanisms of visual encoding, we quantified their waveform characteristics, including latency, peak amplitude, peak time, trough time, and peak-to-trough amplitude. We used these SREP characteristics in subsequent analysis of correlation between SREP peak amplitude and saccade eccentricity, and in machine learning models to map their relationship with successful visual encoding. We calculated SREP characteristics at both averaged and single-trial levels. Here, averaged SREPs refer to SREPs across all saccades, trials, and sessions; and single-trial SREP refers to the LFP trace following a single saccade.

For each electrode contact and cluster identified by the hierarchical clustering procedure described above, we computed the average SREP and quantified its characteristics. We used MATLAB's *findchangepts* function on the first derivative of the averaged SREPs to identify two change points. This function partitions the time series into segments whose means differ maximally, thus detecting points of abrupt change in slope. We defined the first change point SREP latency, corresponding to the earliest statistical deviation from baseline (**Figure 3A**). Within a ±10 ms window around the second change point, we identified the peak time as the point of maximal normalized amplitude for positive-polarity clusters, or minimal amplitude for negative-polarity clusters. We defined trough time as the time point within a 10-ms window around latency at which the normalized amplitude was minimal for positive clusters or maximal for negative clusters.

For each single-trial SREP, we centered the ±10 ms latency search window on the latency of the averaged SREP from its assigned cluster. We applied MATLAB's *findchangepts* function to the first derivative of the single-trial LFP to detect one change point. We defined this change point as the single-trial SREP latency. Similarly, we defined peak amplitude search windows centered on the averaged SREP peak time from the associated cluster. We identified single-trial SREP peak amplitude as the maximum (for positive clusters) or minimum (for negative clusters) amplitude within this window. We identified the single-trial SREP trough between the latency and the peak time, defined as the minimal (for positive clusters) or maximal (for negative clusters) amplitude in this interval. To verify that the hierarchical clustering procedure reliably separated SREP waveforms of two polarities, we quantified the similarity between each single-trial SREP and its cluster-level averaged waveform using Spearman's correlation coefficient. For subsequent analysis, unless specified otherwise, we included single-trial SREPs whose waveforms were similar to the averaged SREP, defined as those with a Pearson correlation coefficient above the median.

Assessing the relationship between SREP characteristics and saccadic properties

To test whether the observed SREPs could be explained by oculomuscular activity, we first examined how SREP amplitude, including peak and peak-to-trough amplitude, varied with Eye-to-Electrode distance. For cluster-level averaged SREP amplitude, we performed linear regression with Eye-to-Electrode distance as the independent variable, using the Python library *statsmodels* (**Figure 3D**). We computed the explained variance, that is, the coefficient of determination (R^2). As a control, we extracted electrooculogram (EOG) signals derived from the differential potential between fronto-polar scalp EEG electrodes (Fp1 and Fp2), and applied the same epoching, clustering, and waveform characterization procedures to EOG signals as for intracranial EPs. We computed R^2 using the same method. We then compared R^2 between intracranial electrode contacts and EOG signals using statistical tests described in the **Quantification and Statistical Analysis** section. We next examined the relationship between saccadic kinematics and SREP characteristics. Specifically, we performed linear regressions between single-trial SREP features, including peak amplitude and peak-to-trough amplitude, and saccade eccentricity. We conducted regression separately for intracranial LFPs and EOG signals, treating eccentricity as the independent variable. As complementary analyses, we calculated the Pearson's correlation coefficient between single-trial SREP amplitude and saccade duration.

To further evaluate whether SREP polarity reflected the movement of extraocular muscles, we examined the relationship between saccade direction and EP polarity for both scalp EOG and

intracranial signals. We extracted the saccade direction associated with each polarity and measured the degree of unimodal bias of the circular distribution of directions using Rayleigh statistics. As a control condition, we computed Rayleigh statistics for all saccade directions pooled across polarities within each electrode contact. We subsequently performed statistical comparisons between monopolar and combined conditions as detailed in the **Quantification and Statistical Analysis** section.

Oscillation detection and quantification

To test if pre-saccadic oscillatory activity influenced SREP polarity, we first identified the dominant oscillation frequency for each electrode contact. We computed the power spectral density (PSD) using Welch's method on continuous re-referenced LFPs. Specifically, we used MATLAB's *pwelch* function using 1-s Hanning windows (2,048 samples) with 50% overlap and a 2-s discrete Fourier transform length (4096 samples). To separate periodic oscillatory components from the aperiodic activity, we fitted a 1/f model to the PSD in log–log space using MATLAB's *polyfit* function. To minimize the influence of 60 Hz line noise, we fitted a linear model to the log-transformed PSD curve within the frequency range of 1-58 Hz. We calculated PSD residues as the pointwise difference between the original PSD curve and the fitted line. To ensure that oscillatory peaks did not bias the aperiodic fit, we excluded frequencies whose residuals were positive and exceeded the median residual, then recalculated the 1/f fit. We defined the dominant oscillation frequency for each contact as the frequency at which the residual PSD curve reached its maximum. To quantify oscillation strength for each electrode contact, we computed the residual PSD value at the dominant oscillation frequency.

To quantify the instantaneous phase and power of the oscillation at the dominant frequency, we applied the Hilbert transform to narrowband-filtered signal segments surrounding each saccade. To ensure that estimates of pre-saccadic oscillatory activity were not affected by the presence of SREP, we segmented LFPs from -3 s before saccade onset up to SREP latency. For each electrode contacts and each encoding task, we estimated the instantaneous oscillation phase and power separately for two clusters of SREPs. To minimize edge artifacts potentially introduced by filtered signals and the Hilbert transform, we symmetrically padded both ends of each segment with 500 ms of data before filtering. We applied a second-order Butterworth band-pass filter centered on the identified dominant oscillation frequency with a half-bandwidth of 3 Hz using MATLAB's butter function. To implement zero-phase digital filtering, we used MATLAB's *filtfilt* function to apply the filter in both the forward and reverse directions. We then used MATLAB's Hilbert function to obtain the analytic signal. To ensure unbiased estimates, we only

derived the instantaneous power and phase from the portions of the analytic signal corresponding to the unpadded original LFPs. Because the oscillation cycle (OC) duration varied with the dominant oscillation frequency, we defined pre-saccadic time points relative to each electrode contact's OC. Our time points of interest included 0.5, 1, 1.5, 2, and 3 OC before the latency.

$$OC = \frac{1}{dominant\ oscillation\ frequency}$$

To study the interplay between SREP and oscillatory activities, we estimated instantaneous oscillation power and phase after SREP latency. For this purpose, we segmented LFPs from -300 to 300 ms relative to the saccade onset and repeated the padding, filtering, and Hilbert-transform procedures. To enable subsequent cross-session and cross-subject comparisons, we z-scored instantaneous power relative to a baseline window, defined as -200 ms to -90 ms relative to saccade onset.

Assessing the relationship between pre-saccadic oscillation phase and SREP polarity

To test whether the oscillatory phase preceding saccades differed between SREP polarities, we first wrapped phase values from [–π, π] to [–2π, 2π] to preserve periodicity and continuity. We then estimated the probability density functions separately for each polarity. To quantify the separation in phase between polarities, we calculated the absolute difference in area under the probability density curve. To determine statistical significance, we generated null distributions by randomly permuting SREP polarity labels (50 iterations per electrode contact) and recalculated the separation metric for each permutation. Observed and null values were compared using two-sided Wilcoxon signed-rank tests (n = 1,970 electrode contact pairs per condition). Effect sizes were quantified using the rank-biserial correlation, which increased progressively from 0.204 to 1.000 across the five pre-saccadic time points, indicating a gradual divergence of oscillatory phase between SREP polarities as saccades approached.

To evaluate whether the pre-saccadic oscillation phase influenced the polarity of SREPs, we quantified phase separation between positive and negative SREP clusters at both the single-contact and contact-averaged levels. Positive and negative SREP clusters refer to sets of SREPs with positive and negative waveform polarities. Analysis in this section used the instantaneous phase derived from LFPs truncated before the SREP latency (see **Oscillation detection and quantification**). Because the phase is circular, direct density estimation on [–π, π] can introduce boundary artifacts, as probability mass at the endpoints (–π and +π) is artificially split. To avoid

this, we extended the phase range by concatenating copies of the phase values shifted by ±2π, preserving continuity across the circular boundary ±π. At the single-contact level, we computed phase histograms for positive and negative polarities using 72 equally spaced bins spanning [–2π, 2π] using the *histogram* function from the *Numpy* Python package. To obtain a continuous phase probability density function (PDF), we applied a Gaussian kernel filter (σ = 1 radian) using the *gaussian_filter1d* function from the SciPy Python package. We normalized each estimated phase PDF such that its integral over the canonical [–π, π] range equaled 1, using the *Simpson* function from the Scipy Python package. We calculated phase separation as the area under the absolute difference between the two normalized PDFs using the same *Simpson* function. At the single-contact level, we computed phase separation for each electrode contact across the five pre-saccadic time points of interest.

To determine statistical significance, we generated null distributions by randomly permuting SREP polarity labels and recalculated the phase-separation metric for each permutation. We conducted 50 permutations per electrode contact. We then compared observed and null separation values at the five pre-saccadic time points of interest, as detailed in the **Quantification and Statistical Analysis** section.

Analysis of oscillatory power and phase synchronization associated with SREPs

To investigate how SREPs interacted with ongoing neural oscillations, we analyzed temporal changes in oscillation power and phase synchronization around saccade times. To quantify inter-trial phase alignment, we computed the inter-trial phase coherence (ITPC) for each polarity and each electrode contact, based on the instantaneous phase of the narrowband analytic signal filtered at each contact's dominant oscillation frequency, which is detailed in the Oscillation detection and quantification section. ITPC is defined as:

$$\mathrm{ITPC}(t,f) = \left| \frac{1}{K} \sum_{\mathrm{k}=1}^{K} e^{i\phi_k(t,f)} \right|,$$

where $\boldsymbol{\phi_k(t,f)}$ denotes the instantaneous phase of saccade $\mathbf{k}$ at time $\boldsymbol{t}$, and $\boldsymbol{f}$ denotes the dominant oscillation frequency. For each electrode contact, we quantified ITPC at time points of interest, including 0.5, 1, 1.5, 2, and 3 OCs before the SREP latency, at latency, peak time, and 0.5 OCs after peak time. To understand whether phase synchronization was accompanied by increased oscillation power, we computed Pearson's correlation between changes in ITPC and changes in instantaneous power relative to 2 OCs before SREP latency at time points of interest, using the *pearsonr* function from the *Scipy* Python library for calculating.

Quantifying within- vs across-contact variability of SREP characteristics

Trial-by-trial differences in recognition memory imply that if SREPs contribute to encoding, within-contact SREP characteristics should vary across saccades and trials. Therefore, before inferring a spatiotemporal pathway from SREP characteristics such as latency and peak amplitude, we established that spatial (e.g., across-contact variability is systematic and not overshadowed by within-contact variability. It is our intention to quantify within-contact versus across-contact (spatial) variability in SREP latency and peak amplitude to determine whether SREPs were anatomically organized. To compute within-contact variability, we first resampled SREP characteristics, including latency and peak amplitude, within each electrode contact to the median saccade count (e.g., 144) across contacts to minimize the effect of sample size on the estimation of variability, thereby ensuring comparability across contacts. We then computed the standard deviation (SD) of SREP latency and peak amplitude for each contact. To estimate across-contact (e.g., spatial) variability with matched sampling, we resampled electrode contacts to equalize the sample number used for within-contact SD estimation. We calculated the standard deviation of latency and peak amplitude from the averaged SREP waveform. We repeated this procedure 100 times and averaged the standard deviation of latency for statistical comparison.

To study whether polarities affected the SREP latency or peak amplitude distribution, at each contact, we computed SREP latency from the trial-averaged waveform separately by polarity and assessed Pearson correlation between the paired latencies or peak amplitudes across contacts. To further validate that the SREP characteristics differences, including latency and peak amplitude differences, between contacts capture the spatial organization of saccadic modulation, we related the between-contact SREP characteristic differences derived from trial-averaged SREPs to the effect size of latency differences estimated from single-trial SREPs. Because the number of contact pairs is large, we randomly sampled 2,000 contact pairs per polarity and computed the Pearson correlation.

Spatial visualization of SREP characteristics

To illustrate the spatial organization of SREPs, we generated both brain region-level and electrode contact-level visualizations. This framework maps outcome values onto a brain model, allowing us to convey findings related to the spatial organization intuitively. We visualized SREP characteristics such as latency and amplitude (**Figure 5B–C**) and other outcomes such as proportion of electrode contact showing early SREP (**Figure 5D**) and feature importance in predicting recognition memory (detailed in **Model Evaluation**) using this framework. For region-level mapping, we assigned each electrode contact an anatomical label using the majority-vote

volumetric approach, detailed in the **Anatomical localization of the electrode** section. We computed the average outcome value, such as the mean SREP latency, across electrode contacts within each region. For each brain region, we reconstructed and rendered three-dimensional surfaces from a reference MRI volume using the PyVista Python package. We mapped regional average values onto the corresponding cortical or subcortical surface meshes and color-coded the surfaces using a sequential colormap (e.g., pale to dark green).

To provide higher-resolution spatial information, we constructed 2D glass-brain heatmaps projected from the 3D brain model. Briefly, we projected electrode coordinates in the common anatomical space onto a 2D plane orthogonal to a vector representing the viewing angle. We discretized this 2D space into grid elements, with each element being a 5×5 square, unless specified otherwise. For each grid element, we computed the local mean outcome values, for example, mean EP latency, across nearby contacts. For each grid cell, we defined nearby contacts as those located within a 10-mm radius. We color-coded each grid element based on the local mean outcome values using a white-to-red colormap. Grid elements without nearby electrode contacts were rendered transparent to truly represent electrode coverage.

Spatial modeling of SREP latency using Gaussian Mixture Models (GMMs)

To examine whether early SREPs were spatially localized, we stratified electrode contacts into percentile bins based on their average SREP latency (e.g., 0–15th, 30–45th, 60–75th percentiles). For each latency percentile bin, we applied 3D GMMs to the spatial coordinates of electrode contacts in the common anatomical space. We implemented GMMs using the *GaussianMixture* function from the *sklearn* Python package, with the number of components (e.g., *n_component*) set to 2 to identify bilateral spatial clusters corresponding to the two hemispheres. We extracted the centroids of identified clusters to determine their hemispheric localization. We quantified the spatial spread as the Euclidean norm of standard deviations along X, Y, and Z axes, estimated by the GMMs. To assess if there is progressive spatial expansion of SREPs across regions, we calculated spatial spread across latency percentile bins and assessed monotonic trends using Spearman's rank correlation coefficient. A significant positive trend would suggest early focal localization of SREP and later widespread involvement.

To quantify the temporal involvement across anatomical regions, we quantified the proportion of electrodes within each region exhibiting SREPs in each latency percentile bin, which was defined as the number of electrode contacts with SREPs in a given latency range divided by the total number of sampled electrodes in that region. To identify regions showing systematic early or late SREPs, we assessed monotonic change using Spearman's rank correlation

coefficient. To examine whether regions showing strong SREPs also exhibited spatial clustering, we repeated the GMM analysis and proportion–percentile correlation analyses described above for SREP peak amplitude.

Saccade-Related Neural Dynamics (SRND) model

SREPs and the preceding oscillatory activity jointly reflect transient neural dynamics that may play a role in visual encoding. The finding that the pre-saccadic oscillation phase influenced SREP polarity, combined with the finding that SREPs were accompanied by oscillation power increase and inter-trial phase coherence, motivated a unified framework, the SRND model, to parameterize and reconstruct the neural mechanism following saccades (**Figure 6A**). For each electrode contact and saccade, we reconstructed the saccade-related neural dynamic $SRND(t)$ as a composite of (1) a sigmoidal SREP waveform and (2) its preceding oscillatory activity. We modeled SREP waveforms $EP(t)$ as sigmoidal rise-and-fall functions parameterized by three temporal features: latency $t_{latency}$, peak time t_{peak}, and peak-to-trough amplitude Amp.

$$EP(t) = OSC\left(t_{latency}\right) + Amp \cdot S\left(t; t_{latency}; t_{peak}\right), \;\; t \leq \; t_{latency}$$

$$S(t) = \begin{cases} \dfrac{1}{1 + exp(-12(\dfrac{t - t_{latency}}{t_{peak} - t_{latency}} - \dfrac{1}{2}))}, & t_{latency} \leq t \leq t_{peak} \\ \dfrac{1}{1 + exp(-12(\dfrac{t - t_{peak}}{t_{end} - t_{peak}} - \dfrac{1}{2}))}, & t_{peak} \leq t \end{cases}$$

We modeled the preceding oscillatory activity $Osc(t)$ as a sine wave parameterized by instantaneous power and phase ϕ of oscillation at the dominant frequency at half OCs before SREP latency.

$$Osc(t) = \sqrt{power}(2\pi f t + \phi)$$

The combined LFP trace $SRND(t)$ is the linear combination of $EP(t)$ and $Osc(t)$. We ensured continuity at SREP latency by adding $Osc\left(t_{latency}\right)$ to $EP(t)$. To truly represent the transient saccadic modulation, we confined the support for time t as [-50, 100] ms relative to saccade onset.

Assessing the relationship between eye movement parameters and encoding performance

To assess whether eye movement parameters were related to visual encoding performance assessed by recognition accuracy on the following day, we compared saccade eccentricity, duration, and direction between remembered and forgotten trials. We quantified the saccade direction dissimilarity as the absolute area between circular kernel density estimates of saccade direction for the two conditions. To evaluate statistical significance, we compared the observed direction dissimilarity with a null distribution generated from 2,000 random permutations of trial labels (remembered vs. forgotten).

Random forest classifier

To test our primary hypothesis that the SRND model predicts visual encoding performance, we trained random forest (RF) classifiers to learn the mapping between SRND parameters and recognition memory outcomes (remembered vs. forgotten), as illustrated in **Figure 6C**. In brief, an RF consists of an ensemble of decision trees, where each tree predicts the class of an observation based on feature thresholds. The final prediction aggregates predicted probabilities across all trees. Each tree is trained on a bootstrap sample of the data. Training a tree involves partitioning the feature space using thresholds that maximize impurity reduction (Gini index). For a node i containing samples from K classes, the Gini index is defined as:

$$Gini\ index\ (i)\ =\ 1\ -\ \sum_{k=1}^{K} p_k^2,$$

where p_k denotes the proportion of samples that belong to class k.

Training and testing data structure

A sample for the RF classifier corresponded to SRND parameters derived from LFPs at a single saccade event from one electrode. We labeled the sample as “remembered” if the image viewed during that saccade’s trial was later recognized by the patient, or “forgotten” if not. Here, a “trial” denotes one image presentation during the encoding task. The input feature vector comprised electrode spatial coordinates along the right-left (x), anterior-posterior (y), and superior-inferior axes (z), and the five SRND parameters. The electrode coordinates (x, y, z) for all participants were expressed in a common anatomical space, as detailed in the **Anatomical localization of the electrode** section. The five SRND parameters included SREP latency, peak-to-trough amplitude, peak time, and instantaneous power and phase of the preceding oscillation at half OCs before SREP latency.

Because multiple saccades, and thus multiple samples across electrode contacts, could occur within a trial, we grouped samples by trial before splitting the dataset into training and testing sets. This choice ensures that all samples from the same trial are assigned exclusively to either training or testing, preventing the model from learning trial-specific idiosyncrasies, such as stimulus content. Therefore, by splitting data at the trial level, we evaluated whether SRND features alone carry information predictive of recognition memory. If the RF classifier successfully predicts recognition from SRND parameters at a specific location, it provides strong evidence that SRNDs reflect neural mechanisms relevant to visual encoding.

Training RF classifiers

Given class imbalance, we implemented balanced RF classifiers using the *imbalanced-learn* and *scikit-learn* Python libraries. A balanced RF classifier draws a bootstrap sample from the minority class (e.g., forgotten) and samples with replacement the same number of samples from the majority class (e.g., remembered). To learn the regional mapping between SRND parameters and visual encoding, while avoiding overfitting to electrode contact-specific locations, we added Gaussian spatial jitter to electrode coordinates in the training data using *normal* function from the *Numpy.random* Python library. We sampled the random *noise* from a normal distribution with standard deviation proportional to the variance of that coordinate across electrodes.

$$spatial\ jitter \sim N(0, 0.1 \times \sqrt{Var_{coordinates}})$$

To reliably evaluate model performance, we implemented 5-fold cross-validation using the *StratifiedKFold* function from the *Scikit-learn* Python library. Unless otherwise specified, we used default hyperparameters of 200 decision trees, a maximum depth of 20, and a minimum leaf size of 1. The results were from this setting unless otherwise noted. To verify that the chosen hyperparameters achieved appropriate model complexity (i.e., neither underfitting nor overfitting), we systematically varied them and reevaluated the model performance (**Figure 6F**).

Model evaluation

Our primary metric for model evaluation was balanced accuracy on held-out test data to account for class imbalance:

$$Balanced\ accuracy\ = \frac{1}{2}(TPR\ +\ TNR),$$

where TPR (true positive rate) is the proportion of correctly classified remembered trials among all remembered trials, and TNR (true negative rate) is the proportion of correctly classified

forgotten trials among all forgotten trials. We also report the area under the receiver operator characteristic curve (ROC AUC) and standard accuracy. ROC AUC summarizes the true-positive vs. false-positive trade-off across thresholds and is resistant to prevalence. Operationally, we constructed a ROC curve by plotting the TPR against the FPR (false positive rate) as the classification threshold varied from 0 to 1.

We evaluated model performance at the location-level, saccade-level, and trial-level. As we observed SREPs across the cerebrum, we reasoned that if the RF model learned local mapping between SRND parameters and visual encoding outcomes, combining predictions across locations within a saccade (e.g., spatial aggregation) should yield a more reliable estimate of the visual encoding outcome than location-level prediction. At the location-level, the model predicted the probability of being remembered and the class of each sample, that is, the set of five SRND parameters at one electrode contact location for a single saccade. For the RF classifier, the predicted probability $\widehat{P_{remember}}(i)$ for a sample i was computed as the mean of predicted probabilities estimated by all decision trees in the ensemble. Tree t's predicted probability for sample i, $p_{i,t}$, equals the proportion of training samples belonging to the "remembered" class in the terminal leaf where that sample falls.

$$\widehat{P_{remember}}(i) = \frac{1}{T}\sum_{t=1}^{T} p_{i,t},$$

where T is the number of trees and $p_{i,t}$ denotes the proportion of "remembered" samples in the terminal leaf reached by sample i in tree t. For each sample, we determined the predicted class label $\widehat{class}_i$ by comparing the predicted probability with 0.5:

$$\widehat{class}_i = \begin{cases} remembered, & if\ \widehat{P_{remember}}(i) \geq 0.5 \\ forgotten, & otherwise. \end{cases}$$

We computed location-level balanced accuracy and ROC AUC directly from these sample-level predictions.

$$ROC\ AUC\ = \int_0^1 TPR(FPR)d(FPR)$$

At the saccade level, we averaged the model's predicted probabilities across all electrode contacts during that saccade to obtain one probability of successful encoding per saccade. We obtained the predicted class $\widehat{class}_i$ similarly and computed balanced accuracy and $ROC\ AUC$ on these saccade-level predicted probabilities and classes. At the trial level, we aggregated predicted

probabilities within a trial to yield a single predicted probability per trial. We then derived the predicted class and computed balanced accuracy and ROC AUC accordingly.

To test whether performance exceeded chance under the null hypothesis of no relationship between SRND features and successful visual encoding, we generated a null distribution by permuting labels. We then retrained and reevaluated the RF models on the permuted data using the same pipeline detailed in **Training and testing data structure**, **Training RF classifiers**, and **Model evaluation sections**. We compared the observed performance metrics, including balanced accuracy and ROC AUC, against their null distribution derived from the permuted data.

To assess whether the model's probabilistic outputs meaningfully aligned with behavioral outcomes, we compared the average predicted probability between remembered and forgotten trials, as a complementary diagnostic of model validity. In addition, we quantified and compared the proportion of predicted remembered vs. forgotten saccades across actual visual encoding outcomes. To assess inter-subject generalizability, we evaluated the model's performance separately for each subject's test data. For each subject, we computed the mean saccade- and contact-level balanced accuracy across CVs. To determine whether observed performance exceeded chance, we generated a subject-specific null distribution by retraining RF models on permuted labels using the identical pipeline. We quantified the percentile rank of each subject's observed performance relative to the null distribution at the individual level (**Figure 6 -Figure Supplement 1E**).

Assessing the reproducibility of learned mappings between SRND and visual encoding outcome

To verify that RF models learned the underlying mapping between SRND parameters and recognition memory performance, we quantified the reproducibility of learned mappings across cross-validations (CVs). We first analyzed the consistency of feature importance across folds. For each fold, the importance of a feature f_j was the mean reduction in Gini index across all decision nodes where the feature was used:

$$FI(f_j) = \frac{1}{T}\sum_{t=1}^{T}\sum_{n\in N_{j,t}}\frac{N_n}{N_t}\,\Delta G(n),$$

where T is the number of trees, $N_{j,t}$ is the set of nodes in tree t that split on feature f_j, N_t is the total number of samples used to train tree t, and N_n is the number of samples reaching node n,

thus, $\frac{N_n}{N_t}$ represents the proportion of samples affected by the split at node n. $\Delta G(n)$ denotes the reduction in Gini impurity index at node n before and after the split:

$$\Delta G(n) = Gini\ index_{\text{parent}}(n) - \left(\frac{N_L}{N_n} Gini\ index_L + \frac{N_R}{N_n} Gini\ index_R\right),$$

where $Gini\ index_{\text{parent}}(n)$ is the Gini impurity index of the parent node n before splitting. $Gini\ index_L$, and $ini\ index_R$ denote the Gini impurity index of the left, and right children's node, respectively, where N_L and N_R denote their respective sample counts, and $N_n = N_L + N_R$.

To complement Gini index-based feature importance measures, we also evaluated the reproducibility of Shapley value–derived feature contributions. Originating from cooperative game theory, the Shapley value for a given sample and a feature represents the feature's marginal contribution to the model output[27]. For instance, for a sample x with model output $f(x)$, the Shapley value $Shapley_j(x)$ for a feature j is the average change in the prediction when adding the feature j to all possible subsets of the other features. (see **Shapley additive explanation** section for detailed computation). For each CV iteration, we computed the mean absolute Shapley value of each feature across training samples, which represents a model-agnostic quantification of each feature's marginal impact on the model output.

We normalized feature importances using the *norm* function from *Numpy.linalg* Python package and ranked them within each CV iteration. We quantified the concordance of feature rankings across folds, for both Gini index-based and Shapley-value derived feature importance, using Kendall's coefficient of concordance (W). Kendall's coefficient of concordance is a rank-based measure of agreement that ranges from 0 (no concordance) to 1 (perfect concordance).

$$W = \frac{12S}{m^2(n^3 - n)}$$

$$S = \sum_{j=1}^{n} \left(R_j - \bar{R}\right)^2$$

m is the number of folds, n is the number of features, R_j is the sum of ranks assigned to feature j across folds, and $\bar{R} = n(m+1)/2$ is the expected rank sum under uniform ranking.

To evaluate the statistical significance of the observed concordance, we generated a null distribution of W values derived from RF models trained on data with permuted class labels. For each permutation, we performed 5-fold CV and obtained a rank vector of feature importance for

the five trained RF models. We then computed W values for 1,000 randomly sampled sets of five permuted-label models, yielding an empirical null distribution for comparison against the observed W.

In addition, we verified that trained RF models across CVs learned similar decision rules. Each decision rule corresponds to a split threshold on a specific feature that maximizes impurity reduction. We reasoned that if independently trained models converged on the same underlying mapping between SRND parameters and memory outcomes, the distributions of these split thresholds should be consistent across folds. For each RF model and each feature, we extracted split thresholds across all nodes and trees. To enable comparability across models, we established a set of reference bins for each feature. Specifically, we computed 100 quantile-based bin edges using the empirical distribution of split thresholds from a reference model (i.e., the model from the first fold). To ensure strictly increasing bin boundaries, we added a small constant (10^{-8}) to the second of any consecutive quantiles with equal values. We then ordered features along the x-axis and constructed a composite split histogram for each RF model, where each bin index represented a unique combination of feature identity and threshold range. In this representation, each histogram summarizes how frequently each feature contributed to splits within specific threshold intervals for a given fold.

To ensure comparability across folds, we normalized each composite split histogram by its total number of splits so that the resulting density distribution summed to one. We quantified the similarity of decision rules between pairs of RF models by computing the area under the absolute difference curve (AUC) between their corresponding kernel density estimates. We obtained kernel density estimates using the *gaussian_kde* function from the *scipy.stat Python* package, with a bandwidth parameter of 0.01.

To assess the statistical significance of the observed decision rule similarity between pairs of RF models, we compared observed similarities from the actual data to a generated null distribution. Specifically, we permuted the class labels 20 times and applied 5-fold cross-validation to each permutation, resulting in a total of 100 RF models trained under random label conditions. From this pool, we randomly selected pairs of RF models and computed the decision rule similarity using the same procedure as for the actual data.

To further demonstrate that RF models convergently learned the mapping between SRND parameters and visual encoding outcome, we examined the consistency of predicted probabilities and classes across models trained on different CV iterations. For each pair of models trained on

actual data but in different folds, we input 200 testing samples from the shared test set, which are observations that were not included in the training data of either model. We then computed Pearson's correlation coefficient between the predicted probabilities of "remembered" for the input testing samples using the *pearsonr* function from the *scipy.stats* Python package. For the same model pairs, we also computed the disagreement ratio as the proportion of test samples for which the predicted classes differed, defined as the. This metric captures divergence in binary decisions rather than probabilistic calibration. To assess the statistical significance of the observed output correlation and disagreement ratio across CV iterations, we generated their corresponding null distributions by randomly pairing RF models trained on permuted labels 200 times and recalculating these two output (dis)agreement metrics.

Logistic regression for linear mapping between SRND parameters and visual encoding outcomes

To determine if a linear relationship could explain the mapping between SRND parameters and visual encoding outcomes, we trained logistic regression models using the same dataset and class labels as those used for training RF models (see **Training RF classifiers** section). Given that the brain's functional organization is spatially nonlinear, we expected the logistic regression models to perform worse than the RF models. Nonetheless, this analysis served as a control to quantify the extent to which a linear model could capture the relationship between SRND features and visual encoding outcomes.

We implemented logistic regression with a logit link function and L2-norm regularization using the *LogisticRegression* class from the *Scikit-learn* Python package. The training and testing data structure followed the same rationale and procedures described in the **Training and Testing Data Structure** and **Training RF Classifiers** sections. To address class imbalance between samples labeled "remembered" and "forgotten", we applied the Synthetic Minority Over-Sampling Technique (SMOTE) using the *SMOTE* function from the *imblearn.over_sampling* package. In brief, SMOTE augments the minority class by generating synthetic samples in feature space[57]. For each minority-class sample, SMOTE generated synthetic examples along the line segments connecting that sample to any of its k minority class nearest neighbors (default k = 5), thereby reducing classifier bias toward the majority class.

As in RF mode training, we introduced spatial jitter to electrode coordinate features to mitigate overfitting to exact electrode coordinates and implemented 5-fold cross-validation (see the **Training RF classifiers** section). After model training, we evaluated the logistic regression model's performance based on the balanced accuracy and ROC AUC calculated from the held-

out testing data. The quantification of these performance metrics and assessment of their statistical significance (i.e., comparing performance metrics to their permutation-based nulls) are detailed in the **Model evaluation** section. To interpret the mapping learned, we extracted the regression coefficients associated with each SRND parameter for each electrode contact location. Positive coefficients indicated that higher feature values increased the likelihood of successful encoding, whereas negative coefficients indicated the opposite. Because logistic regression assumes linearity and monotonic effects, caution is warranted when interpreting coefficients for circular variables, that is, the phase of preceding oscillatory activity.

Shapley additive explanation

To interpret how specific SRND parameters contributed to the random forest (RF) model's predictions of visual encoding outcomes, we employed the Shapley additive explanation framework, a model-agnostic approach for explaining machine learning models[27]. Below, we provide the mathematical definitions and properties of Shapley values as formulated initially by Lloyd Shapley.

For a given sample x_i and feature j, the Shapley value $Shapley_{i,j}$ quantifies the *expected* marginal contribution of feature j to the model's predicted probability of successful visual encoding. The marginal contribution of a feature is defined as the *expected* change in the model's prediction when that feature is included compared with when it is excluded, given a specific subset of other features $S \subseteq F \setminus \{j\}$. Here, F is the full feature set, and S is a subset of F excluding feature j. Operationally, we computed the Shapley value by averaging marginal contributions across all possible subsets S, weighted by the probability of each subset that would occur if we randomly ordered all features. Specifically, a subset S of size $|S|$ would occur with probability $\frac{|S|!(|F|-|S|-1)!}{|F|!}$, which is the number of ways features could be ordered such that the features in S appear before feature j, and the remaining features in $F \setminus S$ appear after j. For example, if $F = (j_1, j_2, j_3, j)$, and $S = \{j_1, j_2\}$, there are $|S|! \times (|F| - |S| - 1)! = 2! \times 1! = 2$ permutations where features in S appear before feature j: $(j_1, j_2, j, j_3)$and (j_2, j_1, j, j_3). Mathematically, Shapley value is defined as:

$$\text{Shapley}_{i,j} = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!\ (|F| - |S| - 1)!}{|F|!} \left[f_{S \cup \{j\}}\left(x_{i,S \cup \{j\}}\right) - f_S\left(x_{i,S}\right) \right]$$

where |S| is the number of features in subset S, and |F| is the total number of features. The term $x_{i,S}$ denotes the subset of feature values from sample x_i corresponding to the feature set $S \subseteq F \setminus \{j\}$, while $x_{i,S\cup\{j\}}$ denotes the same sample but with feature j additionally included. $f_S(x_{i,S})$ represents the expected model prediction when only features in S are known.

$$f_S(x_{i,S}) = E_{x_{F\setminus S}}[f([x_{i,S}, x_{i,F\setminus S}])]$$
$$= \int f([x_{i,S}, x_{i,F\setminus S}])p(x_{F\setminus S}|x_{i,S})dx_{F\setminus S},$$

where $x_{F\setminus S}$ represents all the other feature values that are not in S, $x_{i,S}$ is the subset of feature values for sample i corresponding to features in S. Therefore, $f([x_{i,S}, x_{i,F\setminus S}])$ is the model output from a complete feature vector where the features in subset S take the actual values from sample i (i.e., $x_{i,S}$), and the remaining features in $F \setminus S$ take some possible values drawn from the distribution $\sim p(x_{F\setminus S}|x_{i,S})$. Thus, $f_{S\cup\{j\}}(x_{i,S\cup\{j\}}) - f_S(x_{i,S})$ represents the marginal contribution of feature j to the model's prediction for x_i. The weighting term $\frac{|S|!\,(|F|-|S|-1)!}{|F|!}$ guarantees additivity, meaning that each sample's prediction can be exactly decomposed into the sum of feature contributions. By summing the Shapley value for a sample x_i across all features, we get:

$$\sum_{j\in F} Shapley_{i,j} = \sum_{j\in F}\sum_{S\subseteq F\setminus j} \frac{|S|!\,(|F|-|S|-1)!}{|F|!}[f_{S\cup j}(x_{i,S\cup j}) - f_S(x_{i,S})]$$
$$= f(x_i) - f_\emptyset \;=\; f(x_i) - E[f(x)]$$

In tree-based models, $f_S(x_{i,S})$ can be expressed as the expectation of the model's output conditioned on x_S.

$$f_S(x_{i,S}) = \sum_{L\in \text{leaves}} P(L|x_{i,S})\, v_L,$$

where $x_{i,S}$ denotes the subset of feature values from sample x_i corresponding to the feature set $S \subseteq F$, v_L denotes the predicted value (e.g., predicted probability of successful visual encoding) at leaf L. $P(L|x_{i,S})$ represents the probability of reaching leaf L given only the observed features $x_{i,S}$.

For a given leaf L, we computed this probability $P(L|x_{i,S})$ using the chain rule of probability applied along the path from the root node to L, that is, the product of conditional probabilities of following each successive branch along this path. For node n splits on a feature in S, the observed

feature value $x_{i,S}$ deterministically specifies which branch the sample follows. The probability of taking the branch toward L is 1 if $x_{i,S}$ satisfies the split condition for that branch, and 0 otherwise. If node n splits on a feature not in S, the conditional probability of following the branch toward L is empirical frequency $r_{n\to L}/r_n$ from the training data, where r_n is the number of training samples reaching node n, and $r_{n\to L}$ is the number that proceed down the branch toward L. In other words, given training data, the tree defines a discrete probability distribution over leaves $P(L \mid x_{i,S})$, the likelihood that a random training sample reaches leaf L when only features in S are known.

$$P(L \mid x_{i,S}) = \prod_{n \in path(L)} \begin{cases} 1, & split\ feature(n) \in S\ AND\ x_{i,S}\ directs\ x_i\ toward\ L \\ 0, & split\ feature(n) \in S\ AND\ x_{i,S}\ down\ a\ different\ branch \\ r_{n\to L}/r_n, & split\ feature(n) \notin S \end{cases}$$

To illustrate, consider a simple decision tree in which the root node (Node 0) splits on feature A with a threshold of 5. The left branch ($A < 5$) leads directly to a single leaf L_1 with an output value v_{L_1}= 0.8, while the right branch ($A \geq 5$) leads to another leaf L_2 with v_{L_2} = 0.2. 60 % of the training sample falls on the left side of the split (L_1) and 40 % on the right (L_2). These proportions define the empirical branch probabilities $r_{\text{left}}/r_{\text{root}} = 0.6$ and $r_{\text{right}}/r_{\text{root}} = 0.4$. The table below summarizes the tree structure.

Node	Split	Left	Right	Training proportion
0	Left if $A < 5$	$L_1(v_{L_1} = 0.8)$	$L_2(v_{L_2} = 0.8)$	$r_{0\to L_1}/r_0 = 0.6$ $r_{0\to L_2}/r_0 = 0.4$

To calculate $Shapley_{i,A}$ for $x_{i,A} = 3$, we have

$$Shapley_{i,A} = \sum_{S \subseteq F \setminus \{A\}} \frac{|S|!\ (|F| - |S| - 1)!}{|F|!} \left[f_{S\cup\{A\}}(x_{i,S\cup\{A\}}) - f_S(x_{i,S})\right].$$

As *F* = *A, S* can only be ∅. *Note that* $|\emptyset| = 0$*, therefore*

$$Shapley_{i,A} = \frac{|A|!\ (|A| - |\emptyset| - 1)!}{|A|!} \left[f_A(x_{i,A}) - f_\emptyset\right]$$

$$= f_A(x_{i,A}) - f_\emptyset$$

$$f_\emptyset = \sum_{L \in leaves} P(L|x_{i,\emptyset}) v_L$$

$= r_{0\to L_1}/r_0 \cdot v_{L_1} + r_{0\to L_2}/r_0 \cdot v_{L_2}$

$= 0.6 \times 0.8 + 0.4 \times 0.2 = 0.56$

$$f_A(x_{i,A}) = \sum_{L\in leaves} P(L|x_{i,A})v_L$$

$$= r_{0\to L_1}/r_0 \cdot v_{L_1} + r_{0\to L_2}/r_0 \cdot v_{L_2}$$

$= 1 \times 0.8 + 0 \times 0.2 = 0.8$

Now consider a deeper decision tree in which the root node (Node 0) splits on feature A at a threshold of 5. Samples with $A < 5$ proceed to the left child node (Node 1), whereas those with $A \geq 5$ terminate at a right leaf L_3 (v_{L_3}= 0.4). At Node 1, the data are further partitioned based on feature B with a threshold of 2: samples with $B < 2$ reach leaf L_1 with v_{L_1}=0.6, and samples with $B \geq 2$ reach leaf L_2 with v_{L_2}= 0.5. At the root node, the left and right branch proportions are $r_{0\to 1}/r_0 = 0.6$ and $r_{0\to L_3}/r_0 = 0.4$ respectively. At Node 1, the split on feature B partitions the data evenly, with $r_{1\to L_1}/r_1 = 0.5$ and $r_{1\to L_2}/r_1 = 0.5$. The table below summarizes the tree structure.

Node	Split	Left	Right	Training proportion
0	Left if $A < 5$	$Node\ 1$	$L_3(v_{L_3} = 0.4)$	$r_{0\to 1}/r_0 = 0.6$ $r_{0\to L_3}/r_0 = 0.4$
1	Left if $B < 2$	$L_1(v_{L_1} = 0.6)$	$L_2(v_{L_2} = 0.5)$	$r_{1\to L_1}/r_1 = 0.5$ $r_{1\to L_2}/r_1 = 0.5$

We illustrate how to calculate $Shapley_{i,A}$ with $x_{i,\{A,B\}} = [1,1]$. As $F = \{A, B\}$, S can be $\emptyset$ or $\{B\}$.

$$Shapley_{i,A} = \sum_{S\subseteq F\setminus\{A\}} \frac{|S|!\ (|F| - |S| - 1)!}{|F|!} \left[f_{S\cup\{A\}}(x_{i,S\cup\{A\}}) - f_S(x_{i,S})\right]$$

$$= \frac{1}{2}\left[f_A(x_{i,A}) - f_\emptyset\right] + \frac{1}{2}\left[f_{\{A,B\}}(x_{i,\{A,B\}}) - f_B(x_{i,B})\right]$$

Similarly, we have

$$f_{\emptyset} = \sum_{L \in leaves} P\left(L \middle| x_{i,\emptyset}\right) v_L$$

$$= P\left(L_1 \middle| x_{i,\emptyset}\right) v_{L_1} + P\left(L_2 \middle| x_{i,\emptyset}\right) v_{L_2} + P\left(L_3 \middle| x_{i,\emptyset}\right) v_{L_3}$$

$$= r_{0\to 1}/r_0 \cdot r_{1\to L_1}/r_0 \cdot v_{L_1} + r_{0\to 1}/r_0 \cdot r_{1\to L_2}/r_0 \cdot v_{L_2} + r_{0\to L_3}/r_0 \cdot v_{L_3}$$

$$= 0.6 \times 0.5 \times 0.6 + 0.6 \times 0.5 \times 0.5 + 0.4 \times 0.4 = 0.49$$

$$f_A(x_{i,A}) = \sum_{L \in leaves} P\left(L \middle| x_{i,A}\right) v_L$$

$$= P\left(L_1 \middle| x_{i,A}\right) v_{L_1} + P\left(L_2 \middle| x_{i,A}\right) v_{L_2} + P\left(L_3 \middle| x_{i,A}\right) v_{L_3}$$

$$= 0.5 \times 0.6 + 0.5 \times 0.5 + 0 \times 0.4 = 0.55$$

$$f_B(x_{i,B}) = \sum_{L \in leaves} P\left(L \middle| x_{i,B}\right) v_L$$

$$= P\left(L_1 \middle| x_{i,B}\right) v_{L_1} + P\left(L_2 \middle| x_{i,B}\right) v_{L_2} + P\left(L_3 \middle| x_{i,B}\right) v_{L_3}$$

$$= 0 \times 0.6 + 0 \times 0.5 + 1 \times 0.4 = 0.4$$

$$f_{\{A,B\}}(x_{i,\{A,B\}}) = \sum_{L \in leaves} P\left(L \middle| x_{i,\{A,B\}}\right) v_L$$

$$= P\left(L_1 \middle| x_{i,\{A,B\}}\right) v_{L_1} + P\left(L_2 \middle| x_{i,\{A,B\}}\right) v_{L_2} + P\left(L_3 \middle| x_{i,\{A,B\}}\right) v_{L_3}$$

$$= 1 \times 0.6 + 0 \times 0.5 + 0 \times 0.4 = 0.6$$

Therefore, $Shapley_{i,A} = \frac{1}{2}\,[0.55 - 0.49] + \frac{1}{2}\,[0.6 - 0.4] = 0.13$

Similarly, we have

$$Shapley_{i,A} = \frac{1}{2}\left[f_B\left(x_{i,B}\right) - f_{\emptyset}\right] + \frac{1}{2}\left[f_{\{A,B\}}\left(x_{i,\{A,B\}}\right) - f_A\left(x_{i,A}\right)\right]$$

$$= \frac{1}{2}\,[0.4 - 0.49] + \frac{1}{2}\,[0.6 - 0.55] = -0.02$$

We have $Shapley_{i,A} + Shapley_{i,B} = f(x_i) - E[f(x)] = 0.6 - 0.49 = 0.11$.

We implemented this analytical solution using *TreeExplainer* function from the *SHAP* Python package. For each input sample, this function recursively traverses all decision paths

through each tree, computing $P(L|x_{i,S})$ for each feature subset S. Given that our previous analyses demonstrated high reproducibility of learned mapping across CV iterations, and that Shapley value computation is computationally intensive, we computed Shapley values for the trained RF model in a randomly selected CV iteration. This analysis allows us to identify feature values that increase or decrease the model's predicted probability of successful encoding, thereby translating the RF classifier into interpretable electrophysiological rules.

To further illustrate the meaning of Shapley values, we compared the feature value-Shapley value plots with partial dependence plots (**Figure 7A–B**). A partial dependence plot provides a global, approximate summary that shows the model's output changes as that feature varies while averaging over the effects of all other features. In contrast, Shapley values capture the local (e.g., specific to an anatomical location), sample-specific contribution of each feature by considering all possible feature interactions. To visualize how multiple features jointly influenced model prediction of successful visual encoding, we systematically varied SRND parameters for each electrode contact. We input the feature combination into the trained RF model to estimate the probability of successful encoding. We then reconstructed waveforms from the SRND parameters and assigned the predicted encoding probability to each waveform (**Figure 7C**).

Shapley value-informed thresholds and directions

To better interpret the RF model's mapping between SRND parameters and visual encoding outcomes, we derived Shapley value–informed thresholds and directions for each feature at each electrode contact location. Conceptually, this procedure mimics the decision process of a random forest and allows us to identify the range of SRND parameters conducive to successful visual encoding. For non-circular features, such as peak-to-trough amplitude, we sorted observations by their feature values. We then identified a change point τ_f that minimized the within-segment variance of Shapley values, partitioning values into encoding-favorable versus unfavorable ranges.

$$\tau_f = \arg\min_t \left[\sum_{x_i<t} \left(Shapley_{i,f} - \overline{s_{<t}}\right)^2 + \sum_{x_i\geq t} \left(Shapley_{i,f} - \overline{s_{\geq t}}\right)^2 \right],$$

where $Shapley_{i,f}$ is the Shapley value for sample x_i and feature f, and $\bar{s} < t$ and $\overline{s_{\geq t}}$ are mean Shapley values on either side of the candidate threshold t.

For the circular feature (i.e., instantaneous phase of preceding oscillations), we identified the angular window $\left[\theta_{1_f}, \theta_{2_f}\right]$ on the unit circle that minimized the variance of Shapley values inside versus outside the window.

$$\left[\theta_{1_f}, \theta_{2_f}\right] = \arg \min_{\theta_1, \theta_2 \in [0,2\pi)} \left(\sum_{i \in W(\theta_1, \theta_2)} \left(Shapley_{i,f} - \overline{s_W}\right)^2 + \sum_{i \notin W(\theta_1, \theta_2)} \left(Shapley_{i,f} - \overline{s_{\neg W}}\right)^2 \right),$$

where $W(\theta_1, \theta_2)$ denotes the set of samples whose feature value lies within the angular interval $[\theta_1, \theta_2]$. $\overline{s_W}$ and $\overline{s_{\neg W}}$ denote mean Shapley values of samples inside and outside the candidate window, respectively.

For each feature, we defined the direction (higher vs. lower or inside vs. outside) associated with successful encoding by comparing the mean Shapley values for observations on either side of the identified threshold, or, for circular features, inside vs. outside the angular window. For non-circular features, if the mean Shapley value on one side was higher, feature values that are higher (or lower) than τ_f contributed to successful encoding. For circular features, we computed the mean Shapley value within and outside the angular window $\left[\theta_{1_f}, \theta_{2_f}\right]$. We defined the Shapley value-informed direction as the phase range yielding the higher mean Shapley value. We then quantified how strongly each observation's feature value aligned with the encoding-favorable direction by computing a recall likelihood $L_{recall}(i)$. For a non-circular feature and a sample x_i,

$$L_{recall}(i) = 0.5 \pm \frac{\left|x_{i,f} - \tau_f\right|}{3\sigma_f},$$

where the sign depends on the feature's direction ("higher" or "lower" associated with successful encoding), and $\sigma_f$is the standard deviation of the feature computed from the training data. For circular features, let the phase be $\phi_i \in [0,2\pi)$ and the window $W_f = \left[\theta_{1,f}, \theta_{2,f}\right]$. For sample x_i with direction as "within",

$$L_{recall}(i) = 0.5 + (2 \times I_W(\phi_i) - 1) \times d_W(\phi_i)/\left(3\sigma_f\right),$$

$$I_W(\phi) = \begin{cases} 1, if\ (\theta_1 \leq \theta_2 \text{ and } \theta_1 \leq \phi \leq \theta_2) \text{ or } (\theta_1 > \theta_2 \text{ and } \phi \geq \theta_1 \text{ or } \phi \leq \theta_2) \\ 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad otherwise, \end{cases}$$

$$\Delta_c(a, b) = \min(|a - b|, 2\pi - |a - b|),$$

$$d_W(\phi) = \min\left(\Delta_c(\phi, \theta_1), \Delta_c(\phi, \theta_2)\right).$$

$d_W(\phi)$ represents the minimum circular distance from an angle ϕ to either boundary of the encoding-favorable phase window. We clipped $L_{recall}(i)$ to the interval [0,1].

To validate that the Shapely value-informed thresholds and directions reflected meaningful SRND-visual encoding mappings, we compared recall likelihood averaged across electrode contacts between remembered and forgotten trials. In addition, we evaluated classification performance based on Shapley value-informed thresholds and compared the performance metrics, including balanced accuracy and ROC AUC, to their surrogate null distribution. Specifically, for each feature, we generated surrogate models by uniformly resampling feature thresholds within a range determined by the empirical minimum and maximum of the observed thresholds, while randomly permuting the direction. We then recompute recall likelihoods based on the surrogate configuration. Based on recall likelihood, for each sample and each feature, we classified it as "remembered" if $L_{recall}(i) \geq 0.5$, and those with $L_{recall}(i) < 0.5$ as forgotten. We then computed balanced accuracy based on class prediction and ROC AUC based on recall likelihood, for electrode contact, saccade, and trial level. Please see **Model evaluation** section for computation details.

Visualization of the SRND-visual encoding relationship

After confirming that the Shapley value-informed thresholds and directions distinguished successful from unsuccessful visual encoding outcomes, we visualized them to illustrate how SRND features contributed to successful visual encoding across anatomical regions (**Figure 7I–M**). For non-circular features, we used two complementary color schemes to represent the Shapley-informed direction, that is, whether higher or lower feature values favored successful encoding. A red colormap (light to dark red) indicated that higher feature values were associated with successful encoding, whereas a blue colormap (light to dark blue) indicated that lower feature values were associated with successful encoding. We used color saturation to reflect the magnitude of the threshold, with more saturated hues corresponding to more extreme thresholds, in other words, higher thresholds for "higher" directions and lower thresholds for "lower" directions.

For circular features (i.e., the instantaneous phase at 0.5 OCs before EP latency), we computed the circular mean of the encoding-favorable phase window $\left[\theta_{1_f}, \theta_{2_f}\right]$ when the direction was "within," or its complementary phase range $\neg\left[\theta_{1_f}, \theta_{2_f}\right]$ when the direction was "outside." We color-coded the circular mean of the encoding-favorable phase using the *twilight* circular colormap from the *Matplotlib* Python Package.

To visualize the SRND-visual encoding relationship at electrode contact resolution, we projected the 3D cortical and subcortical surfaces, along with electrode coordinates, onto a 2D plane and discretized this space into a 5 mm × 5 mm grid, as detailed in the **Spatial visualization of SREP characteristics** section. We assigned each electrode contact's mean absolute Shapley value or Shapley value-informed threshold to grid cells within 10 mm. Grid cells with no assigned values were rendered transparent. The resulting heatmaps depict the regional gradient of model contributions, with color intensity encoding the mean absolute Shapley values. Background anatomical surfaces were overlaid for spatial reference and color-coded by anatomical region without encoding Shapley values.

RF regression for predicting saccade direction

To test whether SRND parameters contain information about saccade direction, we split the training and testing datasets at the saccade level and trained RF regressors. The input feature set was identical to that used for visual encoding outcome prediction. Training RF regressors largely followed the same framework described in the **Training RF classifiers** section, including adding Gaussian spatial jitter to electrode contact location and using 5-fold cross-validation. In addition, we verified that the RF regression model was neither overfitting nor underfitting by systematically varying key hyperparameters (**Figure 8-Figure Supplement 1A**). The differences, compared to training an RF classifier to learn the SRND-visual encoding mapping, were the target variable, objective function, and model output. Because saccade direction is a circular variable, its numerical representation in radians is discontinuous at the $\{0, 2\pi\}$ boundary. For instance, 0 rad and 2π rad correspond to the same direction but differ numerically by 2π. To allow the RF regression model to optimize the loss function in Euclidean space and avoid treating saccade direction as a linear quantity, we transformed each saccade's direction θ into a two-dimensional direction vector in Cartesian space:

$$y_i = [\sin(\theta_i),\ \cos(\theta_i)].$$

This maps angular saccade directions to continuous coordinates in Euclidean space, where opposite directions remain far apart in vector space, while equivalent directions are identical, preserving angular continuity. After transformation, we trained RF regression models using a standard mean squared error (MSE) objective function. We implement the RF regression model using *RandomForestRegressor* from the *Scikit-learn* Python package. In brief, each decision tree learned feature-split thresholds that minimized MSE over the two-dimensional direction vectors. For each sample, the trained RF model routes it to a terminal leaf in every individual decision tree. The predicted output for that sample from a single tree is the mean

direction vector computed across all training samples that reached the same terminal leaf. The final ensemble prediction is the average predicted direction vector across all trees $\widehat{y_i}$. We reconstructed the predicted saccade direction $\hat{\theta}_i$ for sample i in angular space using the two-argument arctangent function:

$$\hat{\theta}_{\mathrm{i}} = \mathrm{atan2}\left(\widehat{y_{i_1}}, \widehat{y_{i_2}}\right).$$

RF regression evaluation

We evaluated RF regression model performance on held-out testing data using both MSE and mean absolute error (MAE). Specifically, we calculated the MSE between the predicted and true direction vectors.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N} ||\widehat{y_i} - y_i||^2,$$

where $||\cdot||$ denotes the Euclidean norm. Lower MSE values indicate better prediction accuracy. In addition, we computed the MAE between the reconstructed angular predictions and actual saccade directions

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} \left|\hat{\theta}_i - \theta_i\right|,$$

where angular differences $\widehat{\theta_i} - \theta_i$ were wrapped into the range $[-\pi, \pi]$ before taking the absolute value. We computed the MSE and MAE at both the location and saccade levels. For saccade-level evaluation, we averaged the predicted direction vectors across all electrodes within each saccade. To assess statistical significance, we generated null distributions for MSE and MAE by repeating the cross-validation training procedure with permuted saccade directions. Specifically, for each of 100 permutations, we randomly shuffled target direction vectors y_i, repeated the RF regression model training procedure, and computed the performance metrics.

Quantifying directional modulation

To confirm that SRND parameters did not encode saccade direction for most electrode contacts, we compared the observed testing MSE for each electrode contact against its null distribution from permutations. We calculated the contact-wise percentile rank of testing MSE relative to its null distribution. We identified direction-informative electrode contacts as those with observed

MSEs below the 5th percentile of their null distribution. To complement the model-based approach, we quantified directional non-uniformity of SRND parameters. For each electrode contact and feature, we estimated the circular kernel density of the feature values as a function of saccade direction. Operationally, we discretized the angular space into 60 equal bins covering $0 - 2\pi$ radians. For each bin centered at angle ϕ_j, we computed the mean feature value of saccades within a $\pm 6°$ window around ϕ_j. For circular features (i.e., instantaneous phase), we computed the circular mean using the *circmean* function from the *scipy.stats* Python package. We then applied von Mises kernel regression, implemented via *vonmises* from the *scipy.stats python package* (concentration parameter κ = 15).

For each feature, we quantified directional non-uniformity D_f as the integral of the absolute difference between the circular kernel density and its mean across the angular domain:

$$D_f = \int_0^{2\pi} \left| p_f(\theta) - \overline{p_f} \right| \, d\theta,$$

where $p_f(\theta)$ denotes the kernel density estimate for feature f, and $\overline{p_f}$ is its circular mean. A higher D_f value indicates greater directional modulation.

To evaluate statistical significance, we generated null distributions of D_f by permuting saccade directions 200 times and recomputing the directional non-uniformity for each permutation. We then computed the percentile rank of directional non-uniformity for each electrode contact and feature relative to the corresponding null distribution. For each feature, we computed the proportion of electrode contacts whose directional non-uniformity is less than the 95th percentile of their relative null distribution, indicating a lack of direction-specific modulation.

Neural networks for identifying saccade direction information in SRND

To validate the observed lack of saccade-direction information in SRND parameters, we implemented several complementary modeling approaches. These analyses aim to demonstrate that the lack of saccade-direction encoding across nonlinear, linear, and binary classification frameworks, and therefore confirmed that SRND primarily reflects cognitive rather than motoric components of saccade processing. Specifically, we trained (1) neural networks with either saccade direction or transformed direction vector as output, (2) a linear model for angular prediction, and (3) a binary RF classifier to distinguish leftward versus rightward saccades. These analyses aimed to detect any nonlinear or linear mapping between SRND parameters and saccade direction.

As a nonlinear control model, we implemented a feedforward neural network consisting of three fully connected hidden layers, each followed by a Rectified Linear Unit activation function using the *PyTorch* Python package. The input layer received the same feature set as the RF models, including electrode spatial coordinates and SRND parameters, described in the **Training and testing data structure** section. The network's output layer contained two units representing the two-dimensional Cartesian saccade direction vector $[\sin(\theta), \cos(\theta)]$. To ensure that predictions represented saccade directions, we required the network to normalize its outputs to the unit circle:

$$\hat{y}_i = \frac{u_i}{|u_i|}, \qquad u_i = f_\theta(x_i),$$

where $f_\theta(x_i)$ denotes the network output for input x_i. We trained this neural network to minimize MSE between predicted and true direction vectors. Consistent with previous machine learning model training and evaluation, we used the 5-fold cross-validation procedure. We trained the network using the Adam optimizer from the *PyTorch* Python package with a learning rate of 1×10^{-3}, a batch size of 256, and up to 50 epochs.

To confirm that the transformation of saccade direction into a two-dimensional Cartesian direction vector did not prevent the neural network from learning the relationship between SRND parameters and saccade direction, we implemented a complementary neural network trained to minimize angular loss directly. This model had the same three-layer architecture and inputs as above, but used a single-unit output layer representing the angular direction in radians. We used a loss function to represent the wrapped angular distance between the predicted direction ($\hat{\theta}$) and true directions (θ):

$$\mathcal{L}_{angular} = 1 - cos(\hat{\theta} - \theta).$$

This formulation penalizes large angular deviations and is continuous at the $0 - 2\pi$ boundary. We evaluated both neural networks at the electrode contact and saccade levels using MSE, computed analogously to the RF regression evaluation (see the **RF regression evaluation** section). For the network with two outputs, we reconstructed the predicted angular direction from the predicted direction vector using the two-argument arctangent function. We then computed the MSE between reconstructed and true saccade directions after wrapping angular differences into the range [−180°, 180°]. To evaluate whether prediction accuracy exceeded chance, we generated null distributions of MSE by training neural networks with the same architecture on data with permuted saccade directions. Specifically, we randomly shuffled target direction vectors

$[\sin(\theta), \cos(\theta)]$ across saccades 100 times and repeated the full cross-validation training and evaluation procedure for each permutation.

Kernel regression by minimizing the von Mises negative log-likelihood

To test whether SRND parameters could linearly explain saccade direction, we implemented a linear regression by minimizing the von Mises negative log-likelihood. The von Mises distribution is the circular analog of a Gaussian distribution for angular data, modeling angular data defined on the unit circle while accounting for the periodicity at the 0–2π boundary. Formally, we modelled the saccade direction θ_i as a circular variable whose mean direction μ_i depends linearly on the SRND features:

$$\mu_i = X_i\beta,$$

where X_i is the feature vector comprising SRND parameters and electrode coordinates, and β is the vector of regression coefficients to be estimated. Under the von Mises assumption, the probability density of observing θ_i is given by

$$p(\theta_i \mid \mu_i, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp[\kappa \cos(\theta_i - \mu_i)],$$

where κ is the concentration parameter which we set to 1. $I_0(\kappa)$ is the modified Bessel function of the first kind of order 0. We estimated the regression parameters by minimizing the von Mises negative log-likelihood:

$$\mathcal{L}(\beta) = -\sum_i \kappa \cos(\theta_i - X_i\beta) + N\log\big(2\pi I_0(\kappa)\big),$$

To fit this model, we used the BFGS optimization algorithm implemented via the *minimize* function from the *scipy.optimize* Python package. We obtained the predicted saccade direction from the model as

$$\hat{\mu}_i = X_i\hat{\beta},$$

$$\hat{\theta}_i = \hat{\mu}_i \; mod \; 2\pi,$$

where $\hat{\beta}$ denotes the estimated model parameters, and mod denotes the modulo operation, which ensures that predicted directions remain within the circular domain $[0,2\pi)$.

Consistently, we evaluated model performance using contact- and saccade-level MSEs between predicted and true saccade directions. We assessed the statistical significance of

prediction accuracy by comparing MSEs to their null distributions generated via 100 permutations of saccade labels.

Binary classification of left vs. right saccades

To parallel the classification framework used for predicting successful visual encoding, we also trained a balanced RF classifier to discriminate between leftward and rightward saccades. We binarized saccade directions as left and right based on their angular position. All procedures for data splitting, feature selection, model training, evaluation, and statistical testing followed those described in the **Random forest classifier**, **Training and testing data structure**, **Training RF classifiers**, and **Model evaluation** sections.

Prediction of subsequent saccade direction

Finally, to determine whether SRND represents neural computations related to the planning of subsequent eye movements, we trained an RF regressor to predict subsequent saccade direction based on SRND parameters at the current saccade. We defined subsequent saccades from the eye-tracking data based on spatiotemporal proximity criteria. Specifically, for each valid saccade, we searched for the next non-fixation saccade occurring within the same trial and within 500 ms of the current saccade offset. We required that the gaze positions at the current saccade offset and the next saccade offset differ by less than 20 % of the screen dimension in both x and y coordinates.

For each valid pair of saccades, we used the direction of the next saccade as the target variable in the RF regression model. Model training and evaluation followed the same procedures as described in the **RF regression for predicting saccade direction** section. We quantified prediction performance using MSE between predicted and true subsequent saccade direction vectors and MAE between predicted and true subsequent saccade directions in radius. We assessed statistical significance by comparing observed performance against null distributions derived from models trained on data with permuted saccade directions.

Quantification and Statistical Analysis

All data were analyzed and plotted with MATLAB and Python. For comparison between two conditions, when assumptions of normality and equal variances were met, we used Student's unpaired t-test. When samples were paired (e.g., within-subject or within-electrode comparisons) and these assumptions were met, we used paired Student's t-tests. In cases where these assumptions are not met, we used the Mann–Whitney U test for comparisons between

independent samples, and the Wilcoxon signed-rank test for paired comparisons. For comparison of the mean direction between two (or more) circular distributions, we used Watson-Williams F test. For comparison of the full circular distributions (e.g., differences in shape), we used Mardia–Watson–Wheeler Test. For testing if the distribution of intertrial correlation differed from baseline, we used Kolmogorov-Smirnov tests.

For linear regression analyses, we assessed statistical significance of regression coefficients using two-sided t-tests. For correlation analyses, we converted Spearman's or Pearson's correlation coefficient to t-statistics and assessed their significance against the Student's t-distribution. For testing if learning models performance exceeded chance, we performed permutation testing, in which data labels were randomly permuted to generate a surrogate null distribution; and we compared true performance metric value to this distribution to obtain a percentile-based p-value.

Statistical significance is defined as NS (non-significant), $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$. We report statistical details in the Results section and figure legends, including the statistical tests used, statistics, effect size, and exact value of n. Effect sizes are reported as: regression coefficients for linear models, Pearson's or Spearman's r for correlation analyses, Cohen's d or rank-biserial correlation for pairwise comparisons, and percentile score relative to the null distribution for permutation tests. When more than 10 comparisons were used to support a conclusion, Bonferroni correction was applied.

References

1. Wurtz, R. H. NEURONAL MECHANISMS OF VISUAL STABILITY. *Vision Res.* **48**, 2070–2089 (2008).

2. Katz, C. N. *et al.* A corollary discharge mediates saccade-related inhibition of single units in mnemonic structures of the human brain. *Curr. Biol.* **32**, 3082-3094.e4 (2022).

3. Jutras, M. J., Fries, P. & Buffalo, E. A. Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. *Proc. Natl. Acad. Sci.* **110**, 13144–13149 (2013).

4. Rajkai, C. *et al.* Transient Cortical Excitation at the Onset of Visual Fixation. *Cereb. Cortex* **18**, 200–209 (2008).

5. Ito, J., Maldonado, P., Singer, W. & Grün, S. Saccade-Related Modulations of Neuronal Excitability Support Synchrony of Visually Elicited Spikes. *Cereb. Cortex* **21**, 2482–2497 (2011).

6. Sobotka, S. & Ringo, J. L. Saccadic eye movements, even in darkness, generate event-related potentials recorded in medial septum and medial temporal cortex. *Brain Res.* **756**, 168–173 (1997).

7. Andrillon, T., Nir, Y., Cirelli, C., Tononi, G. & Fried, I. Single-neuron activity and eye movements during human REM sleep and awake vision. *Nat. Commun.* **6**, 7884 (2015).

8. Tan, G., Xu, K., Liu, J. & Liu, H. A Trend on Autism Spectrum Disorder Research: Eye Tracking-EEG Correlative Analytics. *IEEE Trans. Cogn. Dev. Syst.* **14**, 1232–1244 (2022).

9. Lins, O. G., Picton, T. W., Berg, P. & Scherg, M. Ocular artifacts in recording EEGs and event-related potentials. II: Source dipoles and source components. *Brain Topogr.* **6**, 65–78 (1993).

10. Kovach, C. K. *et al.* Manifestation of ocular-muscle EMG contamination in human intracranial recordings. *NeuroImage* **54**, 213–233 (2011).

11. Crapse, T. B. & Sommer, M. A. Corollary discharge circuits in the primate brain. *Curr. Opin. Neurobiol.* **18**, 552–557 (2008).

12. Ali, M. A., Lischka, K., Preuss, S. J., Trivedi, C. A. & Bollmann, J. H. A synaptic corollary discharge signal suppresses midbrain visual processing during saccade-like locomotion. *Nat. Commun.* **14**, 7592 (2023).

13. Wurtz, R. H. Corollary Discharge Contributions to Perceptual Continuity Across Saccades. *Annu. Rev. Vis. Sci.* **4**, 215–237 (2018).

14. Wachowiak, M. All in a Sniff: Olfaction as a Model for Active Sensing. *Neuron* **71**, 962–973 (2011).

15. Hoffman, K. L. *et al.* Saccades during visual exploration align hippocampal 3–8 Hz rhythms in human and non-human primates. *Front. Syst. Neurosci.* **7**, (2013).

16. Leszczynski, M. *et al.* Saccadic modulation of neural excitability in auditory areas of the neocortex. *Curr. Biol.* **33**, 1185-1195.e6 (2023).

17. Renninger, L. W., Verghese, P. & Coughlan, J. Where to look next? Eye movements reduce local uncertainty. *J. Vis.* **7**, 6 (2007).

18. Hughes, A. M., Whitten, T. A., Caplan, J. B. & Dickson, C. T. BOSC: a better oscillation detection method, extracts both sustained and transient rhythms from rat hippocampal recordings. *Hippocampus* **22**, 1417–1428 (2012).

19. Donoghue, T. *et al.* Parameterizing neural power spectra into periodic and aperiodic components. *Nat. Neurosci.* **23**, 1655–1665 (2020).

20. Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N. & Wolpaw, J. R. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.* **51**, 1034–1043 (2004).

21. Molnar, C. *Interpretable Machine Learning: A Guide For Making Black Box Models Explainable*. (Christoph Molnar).

22. Nourmohammadi, A. *et al.* Passive functional mapping of receptive language cortex during general anesthesia using electrocorticography. *Clin. Neurophysiol.* **147**, 31–44 (2023).

23. Pan, Y., Popov, T., Frisson, S. & Jensen, O. Saccades are locked to the phase of alpha oscillations during natural reading. *PLOS Biol.* **21**, e3001968 (2023).

24. Leszczynski, M. & Schroeder, C. E. The Role of Neuronal Oscillations in Visual Active Sensing. *Front. Integr. Neurosci.* **13**, 32 (2019).

25. Jutras, M. J. & Buffalo, E. A. Synchronous Neural Activity and Memory Formation. *Curr. Opin. Neurobiol.* **20**, 150–155 (2010).

26. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).

27. Kuhn, H. W. Classics in Game Theory. 1–328 (2020).

28. Verhagen, J. V., Wesson, D. W., Netoff, T. I., White, J. A. & Wachowiak, M. Sniffing controls an adaptive filter of sensory input to the olfactory bulb. *Nat. Neurosci.* **10**, 631–639 (2007).

29. Barczak, A. *et al.* Dynamic Modulation of Cortical Excitability during Visual Active Sensing. *Cell Rep.* **27**, 3447-3459.e3 (2019).

30. Lahner, B., Mohsenzadeh, Y., Mullin, C. & Oliva, A. Visual perception of highly memorable images is mediated by a distributed network of ventral visual regions that enable a late memorability response. *PLOS Biol.* **22**, e3002564 (2024).

31. Wang, Y., Cao, R. & Wang, S. Encoding of Visual Objects in the Human Medial Temporal Lobe. *J. Neurosci.* **44**, (2024).

32. Doucet, G., Gulli, R. A., Corrigan, B. W., Duong, L. R. & Martinez-Trujillo, J. C. Modulation of local field potentials and neuronal activity in primate hippocampus during saccades. *Hippocampus* **30**, 192–209 (2020).

33. Sander, V., Soper, B. & Everling, S. Nonhuman primate event-related potentials associated with pro- and anti-saccades. *NeuroImage* **49**, 1650–1658 (2010).

34. Staudigl, T., Minxha, J., Mamelak, A. N., Gothard, K. M. & Rutishauser, U. Saccade-related neural communication in the human medial temporal lobe is modulated by the social relevance of stimuli. *Sci. Adv.* **8**, eabl6037.

35. Staudigl, T., Hartl, E., Noachtar, S., Doeller, C. F. & Jensen, O. Saccades are phase-locked to alpha oscillations in the occipital and medial temporal lobe during successful memory encoding. *PLOS Biol.* **15**, e2003404 (2017).

36. Bremmer, F., Kubischik, M., Hoffmann, K.-P. & Krekelberg, B. Neural Dynamics of Saccadic Suppression. *J. Neurosci.* **29**, 12374–12383 (2009).

37. Golan, T. *et al.* Increasing suppression of saccade-related transients along the human visual hierarchy. *eLife* **6**, e27819 (2017).

38. Leszczynski, M. *et al.* Neural activity in the human anterior thalamus during natural vision. *Sci. Rep.* **11**, 17480 (2021).

39. Martin, C. B. & Barense, M. D. Perception and Memory in the Ventral Visual Stream and Medial Temporal Lobe. *Annu. Rev. Vis. Sci.* **9**, 409–434 (2023).

40. DiCarlo, J. J., Zoccolan, D. & Rust, N. C. How Does the Brain Solve Visual Object Recognition? *Neuron* **73**, 415–434 (2012).

41. Prabhakaran, V., Narayanan, K., Zhao, Z. & Gabrieli, J. D. E. Integration of diverse information in working memory within the frontal lobe. *Nat. Neurosci.* **3**, 85–90 (2000).

42. Yang, S. C.-H., Lengyel, M. & Wolpert, D. M. Active sensing in the categorization of visual patterns. *eLife* **5**, e12215.

43. Lee, D. & Malpeli, J. G. Effects of saccades on the activity of neurons in the cat lateral geniculate nucleus. *J. Neurophysiol.* **79**, 922–936 (1998).

44. He, B. J. Scale-free brain activity: past, present, and future. *Trends Cogn. Sci.* **18**, 480–487 (2014).

45. Purpura, K. P., Kalik, S. F. & Schiff, N. D. Analysis of Perisaccadic Field Potentials in the Occipitotemporal Pathway During Active Vision. *J. Neurophysiol.* **90**, 3455–3478 (2003).

46. Kang, I., Talluri, B. C., Yates, J. L., Niell, C. M. & Nienborg, H. Is the impact of spontaneous movements on early visual cortex species specific? *Trends Neurosci.* **48**, 7–21 (2025).

47. Barczak, A. *et al.* Dynamic Modulation of Cortical Excitability during Visual Active Sensing. *Cell Rep.* **27**, 3447-3459.e3 (2019).

48. Buzsáki, G., Anastassiou, C. A. & Koch, C. The origin of extracellular fields and currents-EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* **13**, 407–420 (2012).

49. Zacksenhouse, M. & Ahissar, E. Temporal Decoding by Phase-Locked Loops: Unique Features of Circuit-Level Implementations and Their Significance for Vibrissal Information Processing. *Neural Comput.* **18**, 1611–1636 (2006).

50. Koay, S. A., Charles, A. S., Thiberge, S. Y., Brody, C. D. & Tank, D. W. Sequential and efficient neural-population coding of complex task information. *Neuron* **110**, 328-349.e11 (2022).

51. Williams, C. C. Not all visual memories are created equal. *Vis. Cogn.* **18**, 201–228 (2010).

52. Sapountzis, P., Paneri, S. & Gregoriou, G. G. Distinct roles of prefrontal and parietal areas in the encoding of attentional priority. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E8755–E8764 (2018).

53. Vries, I. E. J. de, Slagter, H. A. & Olivers, C. N. L. Oscillatory Control over Representational States in Working Memory. *Trends Cogn. Sci.* **24**, 150–162 (2020).

54. Backus, A. R., Schoffelen, J.-M., Szebényi, S., Hanslmayr, S. & Doeller, C. F. Hippocampal-Prefrontal Theta Oscillations Support Memory Integration. *Curr. Biol.* **26**, 450–457 (2016).

55. Inman, C. S. *et al.* Direct electrical stimulation of the amygdala enhances declarative memory in humans. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 98–103 (2018).

56. König, S. D. & Buffalo, E. A. A nonparametric method for detecting fixations and saccades using cluster analysis: Removing the need for arbitrary thresholds. *J. Neurosci. Methods* **227**, 121–131 (2014).

57. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).

Supplementary materials

Supplementary Table 1 Data Summary

ID	Experiment duration (day)	Number of study session	Number of test session	Total number of valid visual encoding trials *	Total number of valid saccades
1	4	4	4	184	1010
2	4	8	8	187	1112
3	2	4	4	91	431
4	4	8	8	196	1447
5	6	12	12	263	1561
6	2	4	4	69	392
7	2	4	4	80	1063
8	2	4	4	30	292

*The criteria for a valid visual encoding trial include 1. Based on the eye-tracking data, subjects maintained their gaze on the monitor for more than half the duration of each trial; 2. No brain stimulation was delivered during this trial.

Supplementary Table 2 Patient Characteristics

ID	Epilepsy diagnosis	Sex	Age range	Handedness	WAIS
1	Focal epilepsy with seizure onset in the left hippocampus	F	36–40	N/A	90
2	Multifocal mesial temporal lobe epilepsy	F	36–40	N/A	N/A
3	Multifocal epilepsy with seizure onset in the hippocampus	F	41–45	R	N/A
4	Focal left hippocampal epilepsy	M	36–40	R	98
5	Focal right hippocampal epilepsy	M	36–40	R	76
6	Focal epilepsy in left supplementary motor area	F	36–40	R	N/A

7	No spontaneous seizures recorded	F	31–35	R	107
8	Focal epilepsy with seizure onset in the right hippocampus and amygdala	F	21–25	R	88

N/A: not available. In the Sex column, 'F' stands for 'female' and 'M' stands for 'male'. In the Handedness column, 'L' stands for 'left' and 'R' stands for 'right'. WAIS: Wechsler Adult Intelligence Scale. Ages are reported in 5-year ranges to protect participant confidentiality.

Supplementary Table 3 Visual encoding behavioral results

		The number of				The number of				
ID	Task ID	Visual encoding trial	Tested visual encoding trial	Remembered tested visual encoding trials	Correct recall rate (%)	Test trials	Novel test trials	Correct recognition of novel test trials	Correct recognition rate of novel test trials	Overall correct recognition rate
	All trials									
1	Encoding1	180	160	110	69%	240	80	47	59%	65%
1	Encoding2	180	160	155	97%	240	80	5	6%	67%
2	Encoding1	180	160	83	52%	240	80	72	90%	65%
2	Encoding2	180	160	126	79%	240	80	54	68%	75%
3	Encoding1	180	160	132	83%	240	80	51	64%	76%
4	Encoding1	180	160	121	76%	240	80	68	85%	79%
4	Encoding2	180	160	99	62%	240	80	64	80%	68%
5	Encoding1	180	160	131	82%	240	80	59	74%	79%
5	Encoding2	180	159	140	88%	240	81	42	52%	76%
5	Encoding3	180	160	148	93%	240	80	38	48%	78%
6	Encoding1	160	80	38	48%	160	80	56	70%	59%
7	Encoding1	160	80	71	89%	160	80	65	81%	85%
8	Encoding1	160	80	47	59%	160	80	57	71%	65%

Valid trials only*										
1	Encoding1	91	73	50	68.5%	-	-	-	-	-
1	Encoding2	93	77	74	96.1%	-	-	-	-	-
2	Encoding1	95	77	36	46.8%	-	-	-	-	-
2	Encoding2	92	74	61	82.4%	-	-	-	-	-
3	Encoding1	91	77	64	83.1%	-	-	-	-	-
4	Encoding1	97	80	62	77.5%	-	-	-	-	-
4	Encoding2	99	79	50	63.3%	-	-	-	-	-
5	Encoding1	98	79	62	78.5%	-	-	-	-	-
5	Encoding2	78	67	58	86.6%	-	-	-	-	-
5	Encoding3	87	75	69	92.0%	-	-	-	-	-
6	Encoding1	69	35	14	40.0%	-	-	-	-	-
7	Encoding1	80	40	35	87.5%	-	-	-	-	-
8	Encoding1	30	18	12	66.7%	-	-	-	-	-

Correct recall rate represents the percentage of remembered visual encoding trials over the tested visual encoding trials. The overall correct recognition rate is calculated as the sum of true positives (old stimuli correctly recalled) and true negatives (new stimuli correctly rejected) divided by the total number of test trials. *The criteria for a valid visual encoding trial include 1. Based on the eye-tracking data, subjects maintained their gaze on the monitor for more than half the duration of each trial; 2. No brain stimulation was delivered during this trial.

Supplementary Table 4 Electrode coverage

ID	Task ID	The total number of implanted electrode contacts	The number of bad contacts*	The number of valid contacts	Average number of valid contacts for each subject
1	Encoding1	225	4	221	220.5
	Encoding2	225	5	220	

2#	Encoding1	229	6	223	224.5
	Encoding2	229	3	226	
3#	Encoding1	170	4	166	166
4#	Encoding1	150	4	146	145.5
	Encoding2	150	5	145	
5#	Encoding1	214	3	211	210.3
	Encoding2	214	4	210	
	Encoding3	214	4	210	
6#	Encoding1	229	10	219	219
7#	Encoding1	190	2	188	188
8	Encoding1	222	3	219	219
Total					1592.8

*Electrodes exhibiting epileptiform activity or artifactual signals (e.g., due to poor contact) were excluded prior to re-referencing. # For a subset of subjects, 2 Ad-yech Behnke Fried Depth electrodes, each with 8-contact micro recording, were implanted. Data from micro wires were excluded if they displayed epileptiform activity or artifactual signals. The number of micro wires is not included in this table.

Appendix A Study protocol

The current study is part of the Amygdala Memory Enhancement trial (NCT05065450).

Electrode placement and patient monitoring

Electrodes are implanted stereotactically to help localize the seizure focus. As a standard means to increase accuracy, the surgery is guided by contrasted MR images, with post-operative reimaging for confirmation. The electrodes from various commercial vendors (Adtech, PMT, Dixi) have platinum contacts for recording intracranial electroencephalogram (iEEG) and for stimulation. An electrode may include macroelectrode contacts only or macro- and micro-electrode contacts. All electrodes are FDA-approved equipment for recording intracranial recording during presurgical evaluation.

Epilepsy monitoring unit (EMU) are wards that are designed for inpatient video EEG monitoring and are staffed by a multidisciplinary team of specialists, including attending physicians, fellows, psychiatrists, EEG technicians, nurses, and information-technology technicians. The EMU contains computers with monitors, one station per room, to visualize the real time recording of EEG and video data. The EMU continuously monitors the daily progress and health of each patient server for secure storage and access. The EMU rotates personnel to monitor and care for each patient daily (24 hours per day) during the patient's stay (on average 2 3 weeks) in the EMU. A patient under presurgical evaluation is discharged from the EMU at the discretion of the treating physicians.

All electrophysiological data will be collected through the EMU's clinical intracranial neural recording system and other devices that have been approved by the FDA to be sold for studies in human subjects or are commonly approved and used at other research institutions.

Reference

1. Sharot, T. & Yonelinas, A. P. Differential time-dependent effects of emotion on recollective experience and memory for contextual information. Cognition106, 538–547 (2008).

Figures for the manuscript titled “Mind’s eye: Saccade-related evoked potentials support visual encoding in humans”

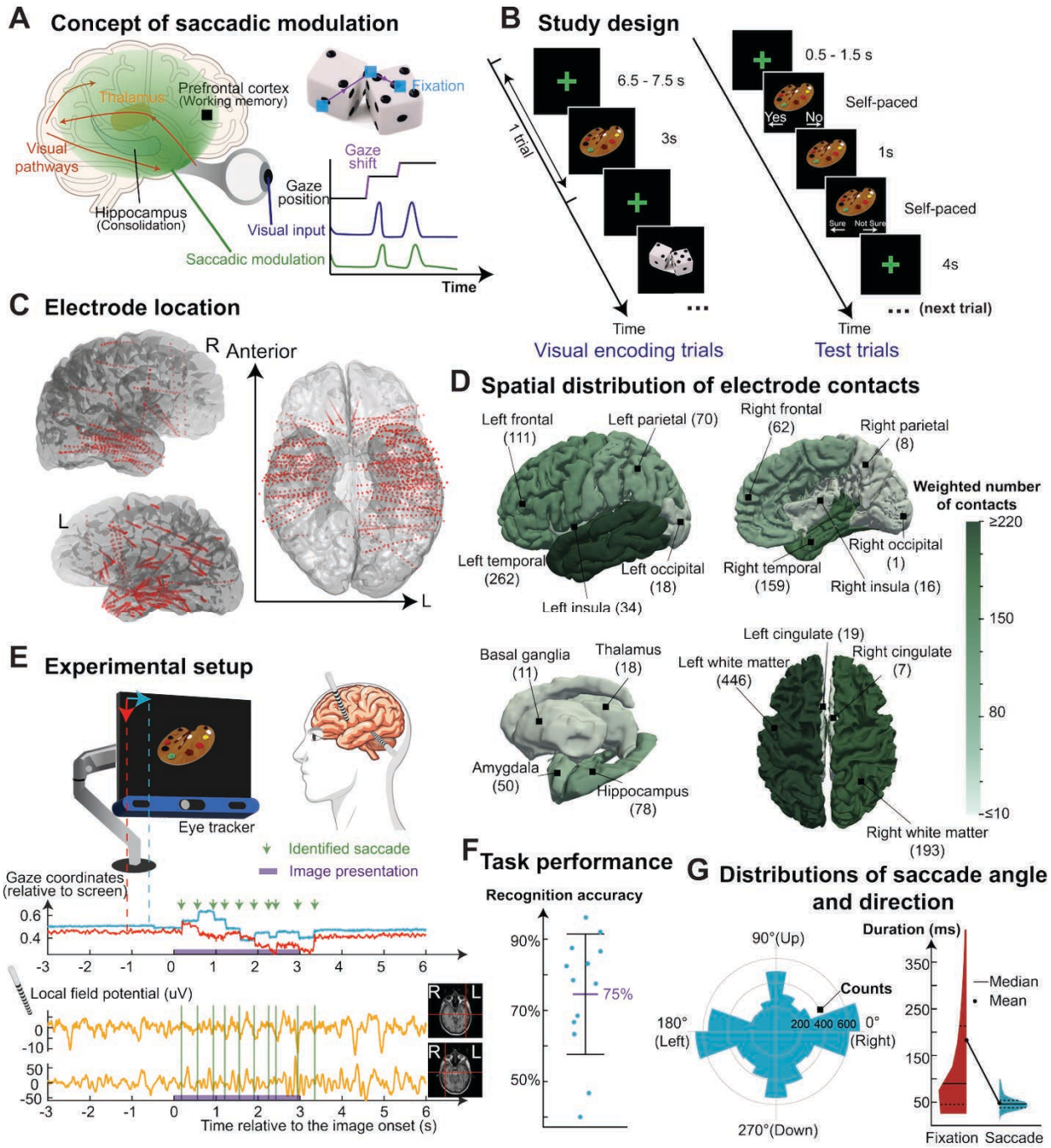


Figure 1 Experimental design for studying saccadic modulation in humans. **(A)** Neural processing of visual information peaks shortly after fixation onset, when visual acuity is at its maximum. Retinal inputs are transmitted to the visual cortex via the lateral

geniculate nucleus. Saccadic modulation refers to changes in neural activities related to eye movements. **(B)** Visual encoding task. Subjects were asked to memorize a series of objects and scenes that were presented for 3 seconds each. Recognition memory was assessed the following day, where subjects used a keyboard to indicate whether they recognized the previously presented images among a series of old and new images. **(C)** Electrode location. We localized implanted stereoelectroencephalography (SEEG) electrodes in subject-specific anatomical space using post-implantation computed tomography (CT) imaging, co-registered to pre-implantation MRI. For visualization purposes, we transformed electrode coordinates into a common anatomical reference space obtained from a reference subject (Subject ID 1). **(D)** Number of electrode contacts per brain region. The common anatomical brain model is divided into four cortical areas (i.e., frontal, temporal, parietal, and occipital cortices), the insula and cingulate cortices, and four subcortical regions (i.e., amygdala, hippocampus, thalamus, and basal ganglia). Color represents the weighted number of electrode contacts per region. When electrode contacts (~2 mm^3 volume) overlapped voxels assigned to multiple brain regions, their contributions were proportionally allocated among these regions (see **Star Methods**). (**E)** Experimental setup. BCI2000 controlled stimulus presentation, recorded local field potentials from SEEG electrodes, and tracked eye movements. We identified saccades onset and offsets based on gaze coordinates using a clustering algorithm (see **Star Methods**). **(F)** Visual encoding task performance. Each data point represents the recognition accuracy for one subject performing an encoding task. We excluded trials in which the eye tracker failed to obtain a valid gaze or in which gaze fell outside the screen boundaries. All subjects confirmed that they performed the task as instructed during a post-experiment debriefing. The average recognition accuracy across tasks and subjects was approximately 75%. (**G**) Distribution of saccade angle and direction.

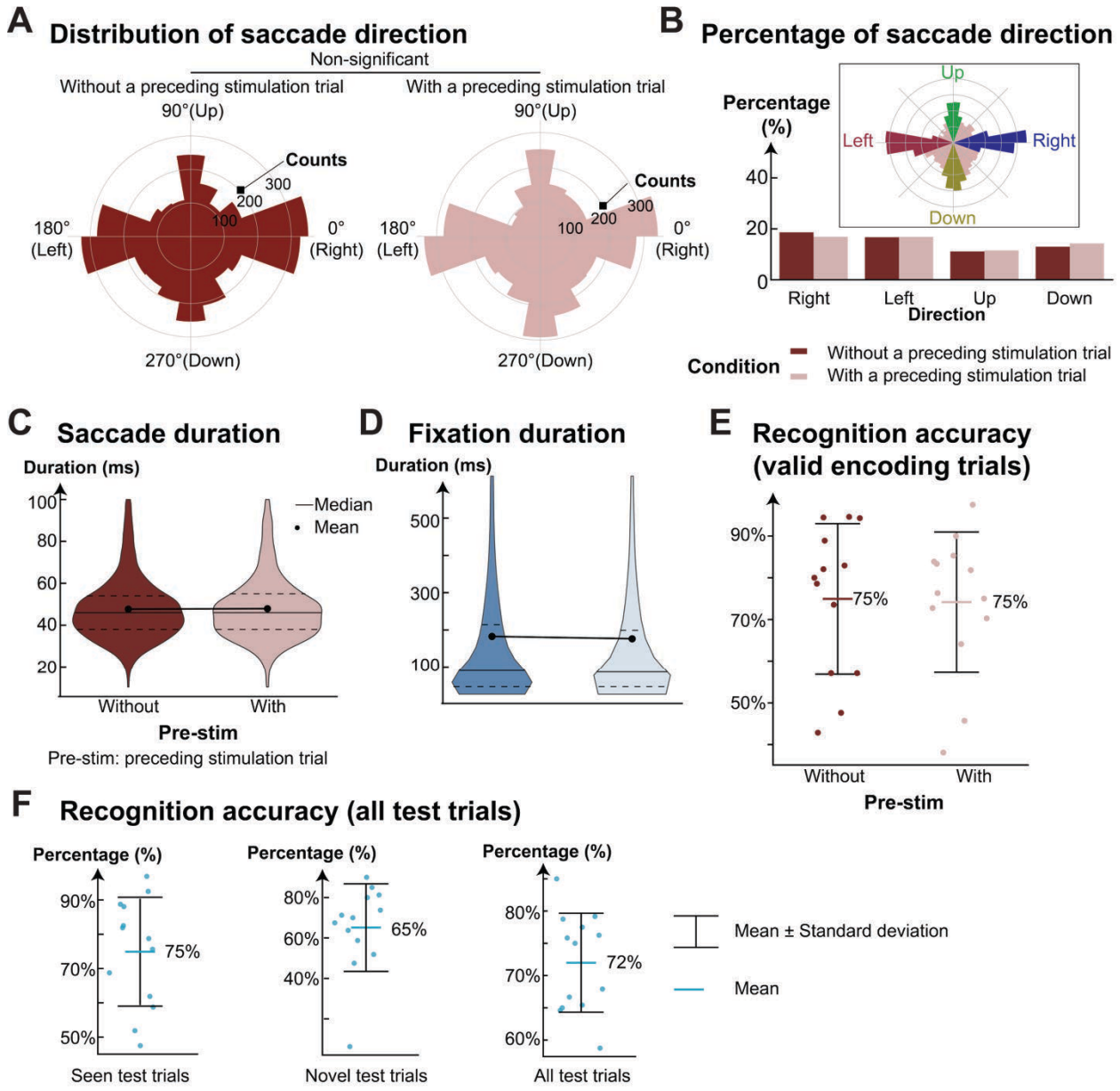


Figure 1 – Figure Supplement 1. Control analysis to rule out any latent effect of electrical stimulation on saccade behavior and recognition memory. This study is part of a clinical trial investigating the effects of electrical stimulation of the basolateral amygdala on visual memory encoding (NCT05065450). Cortical stimulation was administered during 50% of encoding trials. For this study, we excluded all stimulation trials. Comparison of saccade behavior and recognition memory performance for trials without and with a preceding stimulation trial did not show latent effects of electrical stimulation. **(A)** Distribution of saccade directions. We found no significant difference in saccade directions for trials without and with a preceding stimulation trial (Watson-Williams test, F-statistics (1, 7290) = 0.023, p = 0.879). **(B)** Percentage of saccades that are categorized as upward, downward, leftward, or rightward. The inset illustrates the

definition of saccade direction: a saccade is classified as “rightward” if its angle lies within ±20° around the horizontal axis (0°). **(C–D)** Saccade and fixation duration for trials without and with a preceding stimulation trial. The effect size comparing durations between the two conditions was trivial (|Cohen’s d| < 0.05). **(E)** Recognition accuracy for trials without and with a preceding stimulation trial (Cohen’s d = -0.044). **(F)** Recognition accuracy for seen (e.g., old) test trials, novel test trials, and all test trials.

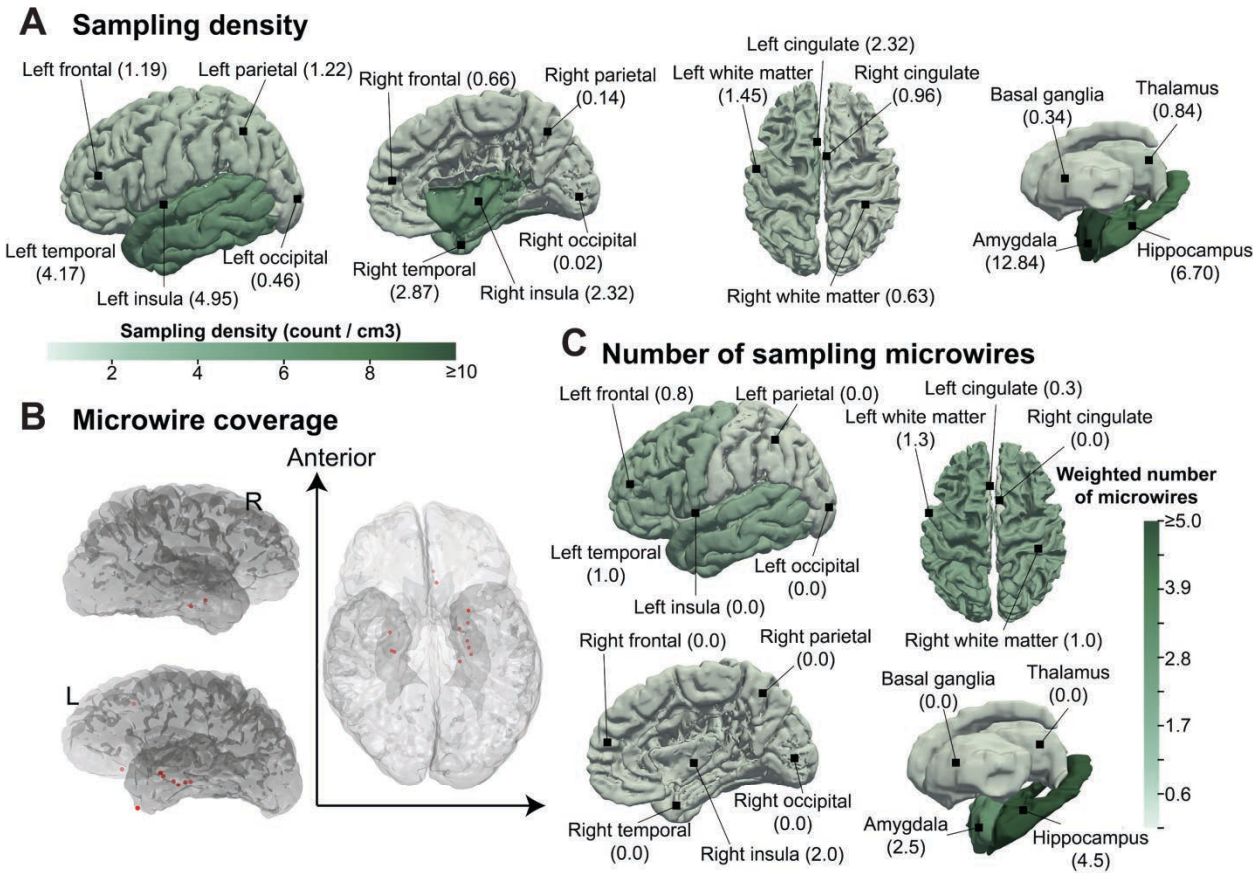


Figure 1 – Figure Supplement 2. Macro electrode sampling density and microwire locations. Six subjects were each implanted with two Ad-Tech Behnke-Fried depth electrodes, each containing eight macro contacts and eight microwires. We refer to the bundle of microwires emerging from the tip of each electrode as a micro-contact location. **(A)** Sampling density for macro-electrode contacts, calculated by dividing the weighted contact count within each brain region by the volume of that region. **(B)** Location of microwires on the brain model of a reference subject. **(C)** Number of microwires in major brain regions.

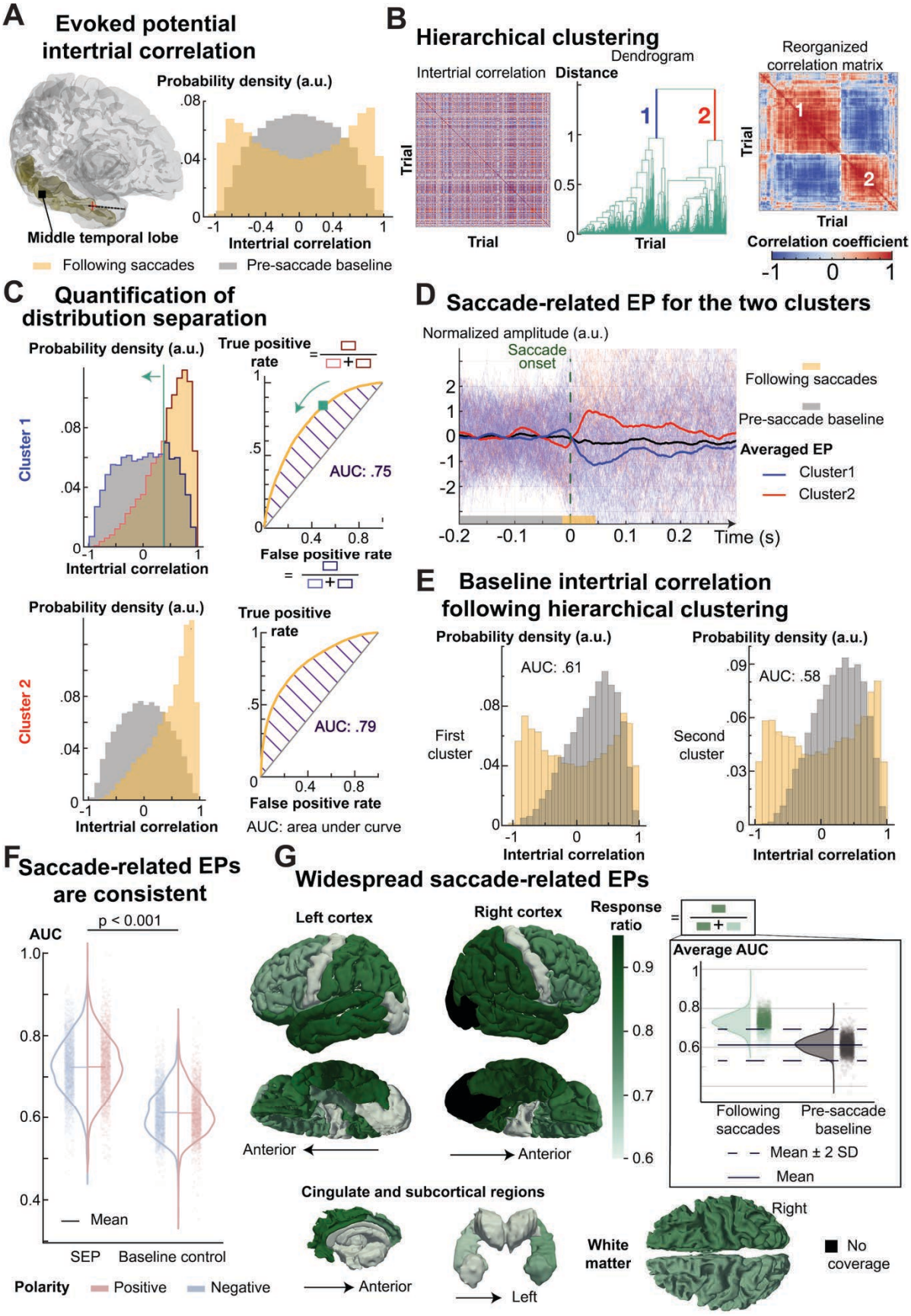

A
Evoked potential intertrial correlation
Probability density (a.u.)
Middle temporal lobe
Intertrial correlation
Following saccades
Pre-saccade baseline
B
Hierarchical clustering
Intertrial correlation
Dendrogram
Distance
Reorganized correlation matrix
Trial
Correlation coefficient
C
Quantification of distribution separation
Cluster 1
Cluster 2
True positive rate
False positive rate
AUC: .75
AUC: .79
AUC: area under curve
D
Saccade-related EP for the two clusters
Normalized amplitude (a.u.)
Saccade onset
Following saccades
Pre-saccade baseline
Averaged EP
Cluster1
Cluster2
Time (s)
E
Baseline intertrial correlation following hierarchical clustering
First cluster
Second cluster
AUC: .61
AUC: .58
F
Saccade-related EPs are consistent
AUC
p < 0.001
Mean
SEP
Baseline control
Polarity
Positive
Negative
G
Widespread saccade-related EPs
Left cortex
Right cortex
Response ratio
Average AUC
Following saccades
Pre-saccade baseline
Mean ± 2 SD
Mean
Anterior
Cingulate and subcortical regions
Left
White matter
Right
No coverage

Figure 2. Evoked potentials following saccades are consistent and widespread across the cerebrum. **(A)** Intertrial correlation of local field potentials (LFPs). To determine the presence of saccade-related evoked potentials (SREPs), we calculated the intertrial correlation of LFPs during the peri-saccadic and the baseline window (i.e., -10 ms to 40 ms, and -200 ms to -10 ms, relative to saccade onsets, respectively). Here, each trial is defined as an LFP segment centered on a saccade event. We calculated the intertrial correlation across all trial combinations using Pearson's correlation. The exemplar result for a representative electrode contact within the right middle temporal lobe shows an intertrial correlation distribution during the peri-saccadic period skewed toward the tails. In contrast, the corresponding baseline intertrial correlation appears normally distributed. **(B)** Hierarchical clustering. The bimodal nature of the peri-saccadic correlation distribution suggests the presence of two EP clusters, with trials within a cluster positively correlated and those between clusters negatively correlated. We used hierarchical clustering to identify these clusters. The dendrogram and the reordered intertrial correlation matrix confirm the presence of two distinct saccade-evoked potentials (shown in D). **(C)** Quantification of the separation between peri-saccadic and baseline correlation distributions. We used a sliding threshold to compute the receiver operator characteristic (ROC) curve. For a set of thresholds from -1 to 1, we calculated the true positive rate (fraction of peri-saccadic probability density above a threshold relative to the combined density) and false positive rate (fraction of peri-saccadic probability density below a threshold relative to the combined density) for each threshold. We used the area under the ROC curve (AUC) to measure separability. **(D)** Evoked potentials averaged across trials in clusters 1 and 2. To extract EP waveforms, we z-scored LFPs before hierarchical clustering. Each line represents the z-scored LFP of one trial. **(E)** Hierarchical clustering on baseline data. This procedure accounts for clustering-induced AUC inflation. **(F)** Peri-saccadic and baseline AUCs for intertrial correlations. AUC following saccades was significantly higher than baseline AUCs (Cohen's d = 1.857, two-tailed paired t-test, n = 2604, t statistic = 81.13, $p < 0.001$), indicating robust saccade-related EP. Each dot represents one electrode contact in one subject during one task. Colors denote the EP polarity based on the sign of the first peak (details in **Figure 3**). **(G)** Response ratios across cerebrum. Represented by the color, the response ratio is defined as the proportion of electrode contacts with AUC in the peri-saccadic condition above the mean + 2 standard deviation (SD) of baseline AUC. Electrode contacts are co-registered to a common reference brain model. The brain was segmented into White Matter (Left, Right), Hippocampus (Left, Right), Amygdala (Left, Right), Basal Ganglia, Thalamus, Orbitofrontal (Left, Right), Inferior Frontal (Left, Right), Middle Frontal (Left, Right), Superior Frontal (Left, Right), Anterior Cingulate, Posterior Cingulate, Motor (Left, Right), Anterior Temporal (Left, Right), Middle Temporal (Left, Right), Superior Temporal (Left, Right), Inferior Temporal (Left, Right), Medial Temporal (Left, Right), Fusiform (Left, Right), Parietal (Left, Right), Occipital (Left, Right), and Insula (Left, Right). We show robust

results of response ratio across cerebrum across various responsiveness thresholds in **Figure 2 – Figure supplement 1**, for example, mean + 3 SD of baseline AUC. We show the spatial distribution of intertrial consistency magnitude in **Figure 2 – Figure Supplement 2**. We show the electrode contact-level intertrial consistency magnitude in **Figure 2 – Figure Supplement 3**.

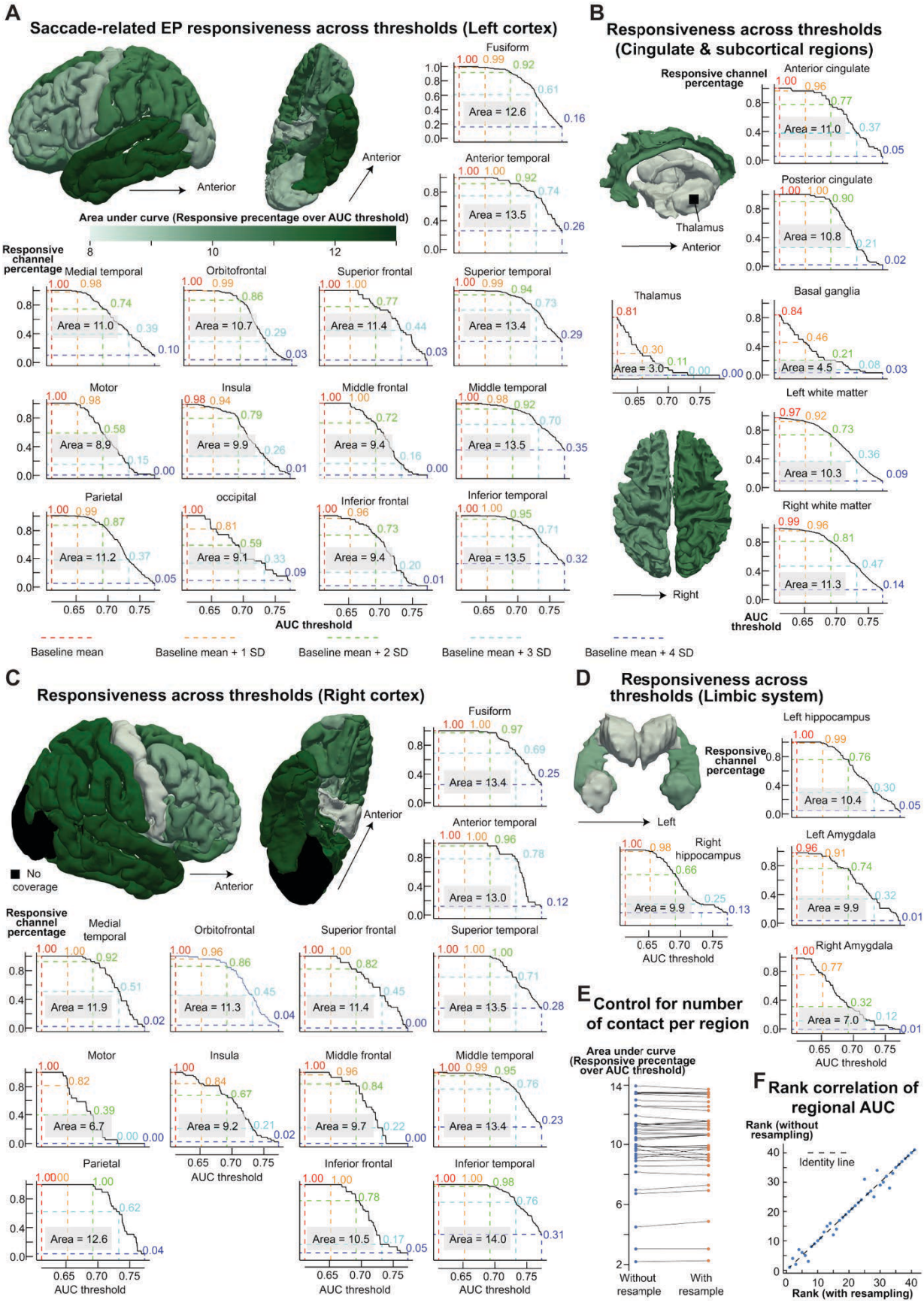

A
Saccade-related EP responsiveness across thresholds (Left cortex)
Anterior
Anterior
Area under curve (Responsive precentage over AUC threshold)
8
10
12
Responsive channel percentage
Fusiform
Area = 12.6
Anterior temporal
Area = 13.5
Medial temporal
Area = 11.0
Orbitofrontal
Area = 10.7
Superior frontal
Area = 11.4
Superior temporal
Area = 13.4
Motor
Area = 8.9
Insula
Area = 9.9
Middle frontal
Area = 9.4
Middle temporal
Area = 13.5
Parietal
Area = 11.2
occipital
Area = 9.1
Inferior frontal
Area = 9.4
Inferior temporal
Area = 13.5
AUC threshold
Baseline mean
Baseline mean + 1 SD
Baseline mean + 2 SD
Baseline mean + 3 SD
Baseline mean + 4 SD
B
Responsiveness across thresholds (Cingulate & subcortical regions)
Responsive channel percentage
Anterior cingulate
Area = 11.0
Thalamus
Anterior
Posterior cingulate
Area = 10.8
Thalamus
Area = 3.0
Basal ganglia
Area = 4.5
Left white matter
Area = 10.3
Right
Right white matter
Area = 11.3
AUC threshold
C
Responsiveness across thresholds (Right cortex)
No coverage
Anterior
Anterior
Responsive channel percentage
Fusiform
Area = 13.4
Anterior temporal
Area = 13.0
Medial temporal
Area = 11.9
Orbitofrontal
Area = 11.3
Superior frontal
Area = 11.4
Superior temporal
Area = 13.5
Motor
Area = 6.7
Insula
Area = 9.2
Middle frontal
Area = 9.7
Middle temporal
Area = 13.4
Parietal
Area = 12.6
Inferior frontal
Area = 10.5
Inferior temporal
Area = 14.0
AUC threshold
D
Responsiveness across thresholds (Limbic system)
Left
Responsive channel percentage
Left hippocampus
Area = 10.4
Right hippocampus
Area = 9.9
Left Amygdala
Area = 9.9
Right Amygdala
Area = 7.0
AUC threshold
E
Control for number of contact per region
Area under curve (Responsive precentage over AUC threshold)
Without resample
With resample
F
Rank correlation of regional AUC
Rank (without resampling)
Identity line
Rank (with resampling)

Figure 2 – Figure Supplement 1. **The percentage of electrodes showing consistent saccade-related EP across responsiveness thresholds**. **(A)** Saccade-related EP consistency in the left cortex. Each panel shows the proportion of electrode contacts whose peri-saccadic AUC is over thresholds. The color-coded dashed line represents the reference threshold, mean, mean + 1 standard deviation (SD), mean + 2 standard deviation (SD), mean + 3 standard deviation (SD), and mean + 4 standard deviation (SD) of baseline AUC. The color of the brain model reflects the area under the curve (AUC) computed across thresholds, representing saccade-related EP consistency. **(B–D)** Saccade-related EP consistency in the right cortex, subcortical regions, and cingulate cortex. **(E)** Areas under the curve of the responsiveness ratio across the AUC thresholds. To control for the number of sampling electrodes in different regions, we resampled electrodes with replacement to equalize sample sizes across regions, then repeated the calculation of the area under the curve. Each dot represents the area under the curve for each region with or without resampling. **(F)** Correlation between AUCs derived from original and resampled data. The strong correlation revealed that the number of sampling electrodes has a trivial effect on the rank of SEP consistency (Spearman's $\rho = 0.987$, $p < 0.001$, $n = 37$, two-tailed t-test).

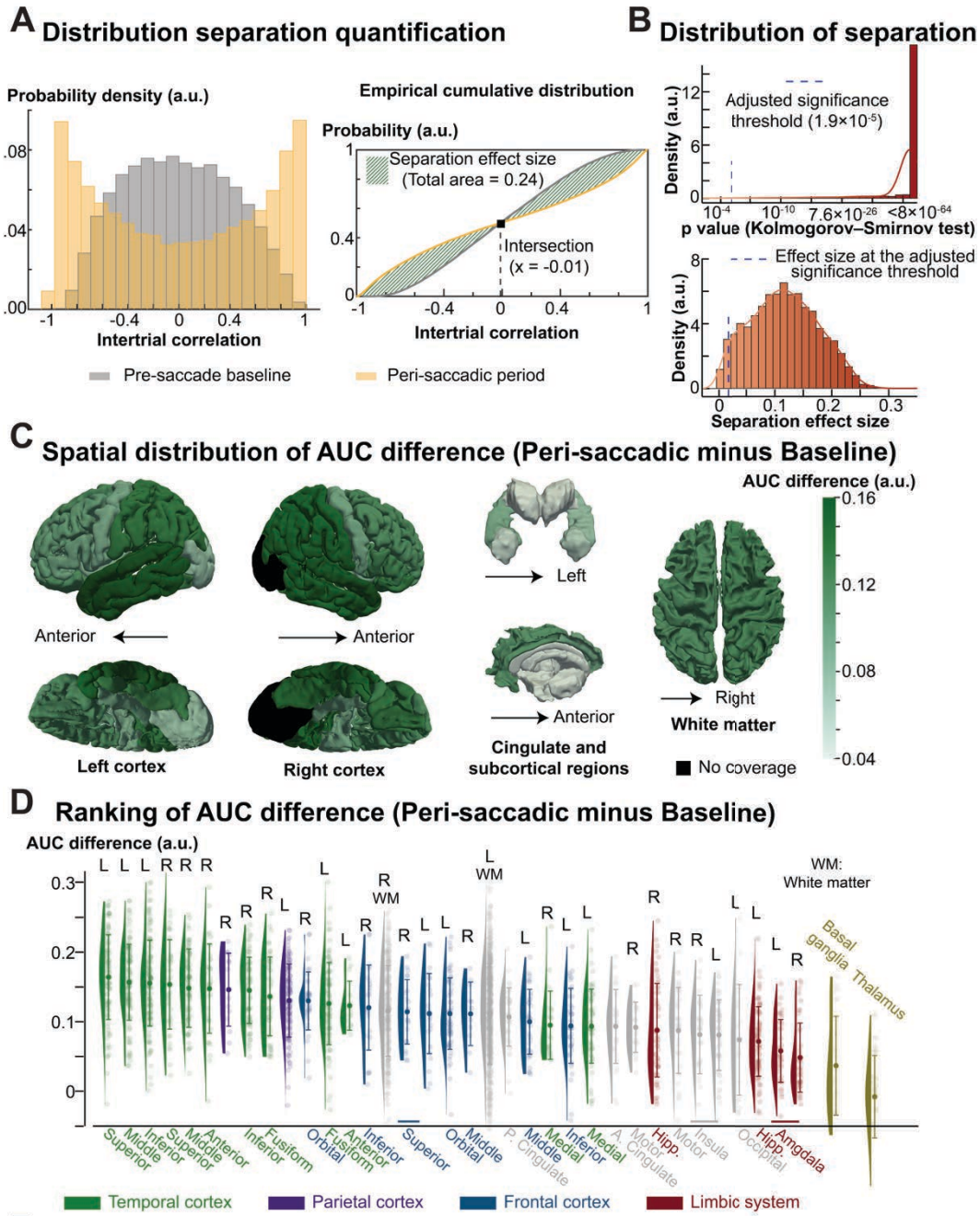

A Distribution separation quantification
Probability density (a.u.)
Intertrial correlation
Pre-saccade baseline
Empirical cumulative distribution
Probability (a.u.)
Separation effect size
(Total area = 0.24)
Intersection
(x = -0.01)
Intertrial correlation
Peri-saccadic period
B Distribution of separation
Adjusted significance threshold (1.9×10-5)
Density (a.u.)
p value (Kolmogorov–Smirnov test)
Effect size at the adjusted significance threshold
Separation effect size
C Spatial distribution of AUC difference (Peri-saccadic minus Baseline)
AUC difference (a.u.)
Anterior
Left
Right
Left cortex
Right cortex
Cingulate and subcortical regions
White matter
No coverage
D Ranking of AUC difference (Peri-saccadic minus Baseline)
AUC difference (a.u.)
WM: White matter
Temporal cortex
Parietal cortex
Frontal cortex
Limbic system
E Spatial distribution of average AUC for positive and negative clusters
Average AUC
Anterior
Left
Right
Left cortex
Right cortex
Cingulate and subcortical regions
White matter

Figure 2 – Figure Supplement 2. Quantifying the regional saccade-related EP consistency based on intertrial correlation distributions. **(A)** Quantification of the separation between intertrial correlation distributions during the baseline and peri-saccadic periods. To statistically assess the presence of robust and bimodal saccade-related EP, we quantified the separation (e.g., deviation from normal distribution toward bimodal distribution) between the intertrial correlation distributions during the peri-saccadic and baseline periods. Specifically, we computed the intersection point of their cumulative distributions and measured the absolute difference in area under the curve (AUC) between them. The striped region illustrates the absolute difference in AUC. **(B)** Distributions of p-value from Kolmogorov-Smirnov tests and the distribution of the separation effect size. The significance threshold was corrected for multiple comparisons. The blue dashed line marks the bimodal deviation corresponding to the significance threshold. The effect size corresponding to the adjusted significance threshold is 0.017. **(C)** The separation between the intertrial correlation distributions during the peri-saccadic and baseline periods across the cerebrum. The color reflects the difference in the area under the curve (computed separately for the two clusters and then averaged) of probability density over intertrial correlation (**Figure 2c**). **(D)** The ranking of the magnitude of SEP consistency (AUC difference). Colors indicate anatomical regions. L: Left; R: Right. **(E)** The magnitude of saccade-related EP consistency across anatomical regions. We quantified the magnitude of EP consistency as the average area under the probability density curve for intertrial correlation following saccades across polarity clusters.

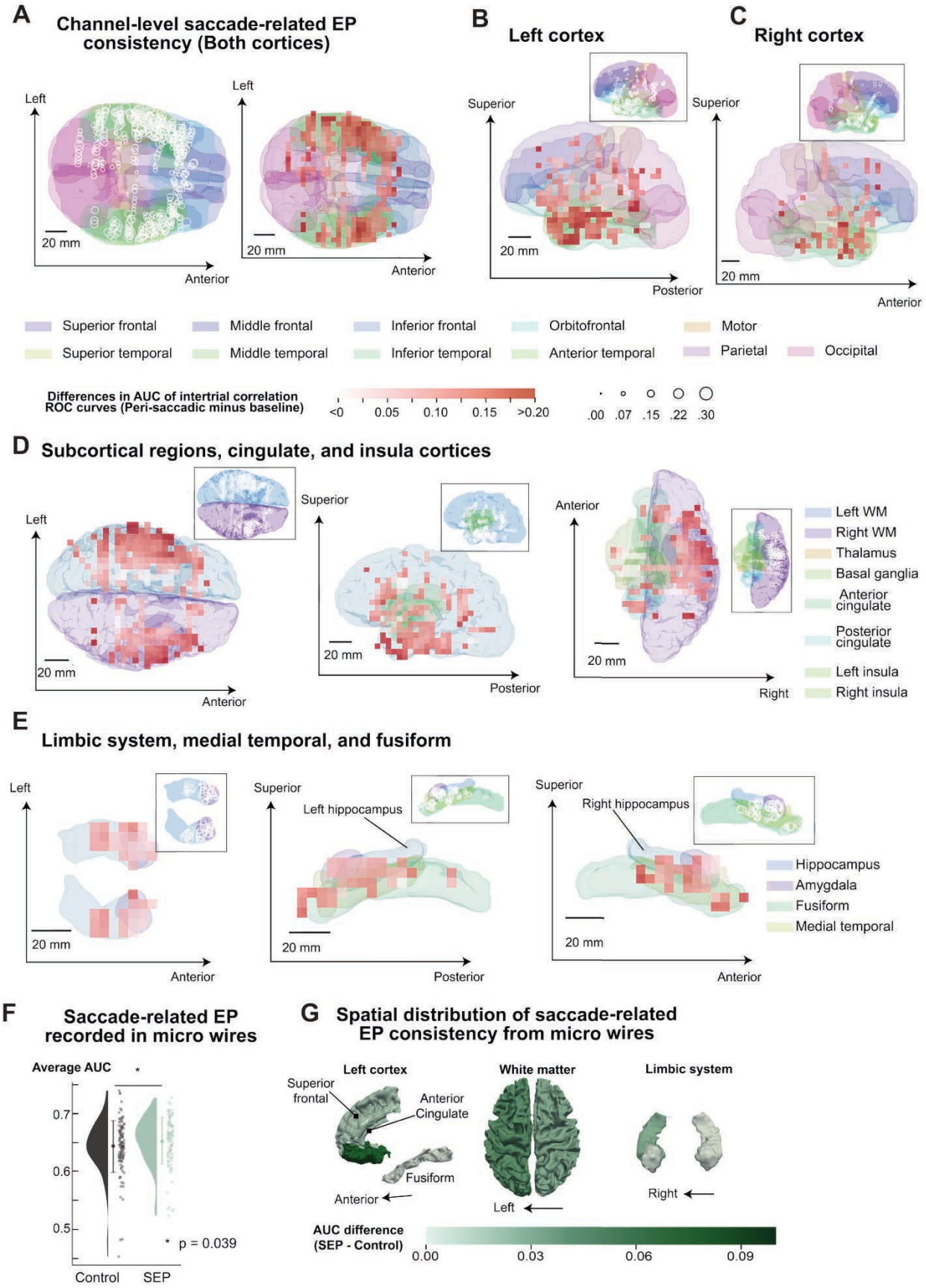

A
Channel-level saccade-related EP consistency (Both cortices)
Left
20 mm
Anterior
Left
20 mm
Anterior
B
Left cortex
Superior
20 mm
Posterior
C
Right cortex
Superior
20 mm
Anterior
Superior frontal
Middle frontal
Inferior frontal
Orbitofrontal
Motor
Superior temporal
Middle temporal
Inferior temporal
Anterior temporal
Parietal
Occipital
Differences in AUC of intertrial correlation ROC curves (Peri-saccadic minus baseline)
<0
0.05
0.10
0.15
>0.20
.00
.07
.15
.22
.30
D
Subcortical regions, cingulate, and insula cortices
Left
20 mm
Anterior
Superior
20 mm
Posterior
Anterior
20 mm
Right
Left WM
Right WM
Thalamus
Basal ganglia
Anterior cingulate
Posterior cingulate
Left insula
Right insula
E
Limbic system, medial temporal, and fusiform
Left
20 mm
Anterior
Superior
Left hippocampus
20 mm
Posterior
Superior
Right hippocampus
20 mm
Anterior
Hippocampus
Amygdala
Fusiform
Medial temporal
F
Saccade-related EP recorded in micro wires
Average AUC
*
0.7
0.6
0.5
* p = 0.039
Control
SEP
G
Spatial distribution of saccade-related EP consistency from micro wires
Left cortex
Superior frontal
Anterior Cingulate
Fusiform
Anterior
White matter
Left
Limbic system
Right
AUC difference (SEP - Control)
0.00
0.03
0.06
0.09

Figure 2 – Figure Supplement 3. Electrode contact-level differences in the area under the curve (AUC) of intertrial correlation receiver operation characteristic (ROC) curves between baseline and peri-saccadic periods. For each electrode contact and each condition (baseline vs peri-saccade), AUCs were calculated separately for the two SREP clusters and then averaged. AUC difference between conditions, the degree of SREP consistency in anatomical regions sampled by each electrode. To visualize spatial patterns and enable comparison of effects across panels (devoid of a 3D perspective), we projected the 3D brain surface and electrode coordinates onto a 2D plane. The overlay heatmap represents the spatial pattern of SREP consistency. To generate the heatmap, we discretized this 2D space into a regular grid with a 5 mm resolution. We computed the local average AUC difference within a 10 mm radius for each grid element. Heatmap elements that have no nearby sampling electrodes are fully transparent. **(A–E)**. Spatial pattern of SREP consistency in cortices and subcortical regions. The heatmap color bar is uniform across panels A–E. **(F)** The consistency of peri-saccadic EPs recorded from micro wires. We calculated the area under the receiver operator characteristic curve from intertrial correlation distributions. We averaged the area under the curve (AUC) for the two clusters. The average peri-saccadic AUCs were higher than the control AUCs (Cohen's $d = 0.227$, two-tailed paired t-test, $n = 115$, t statistic $= 2.088$, $p < 0.039$). **(G)** Spatial distribution of EP consistency across the cerebrum. Brain regions located within 5 mm of the microwire electrodes are shown, which include Left Hippocampus, Left Amygdala, Left White Matter, Right Hippocampus, Right Amygdala, Right White Matter, Left Fusiform, Left Superior Frontal, and Anterior Cingulate.

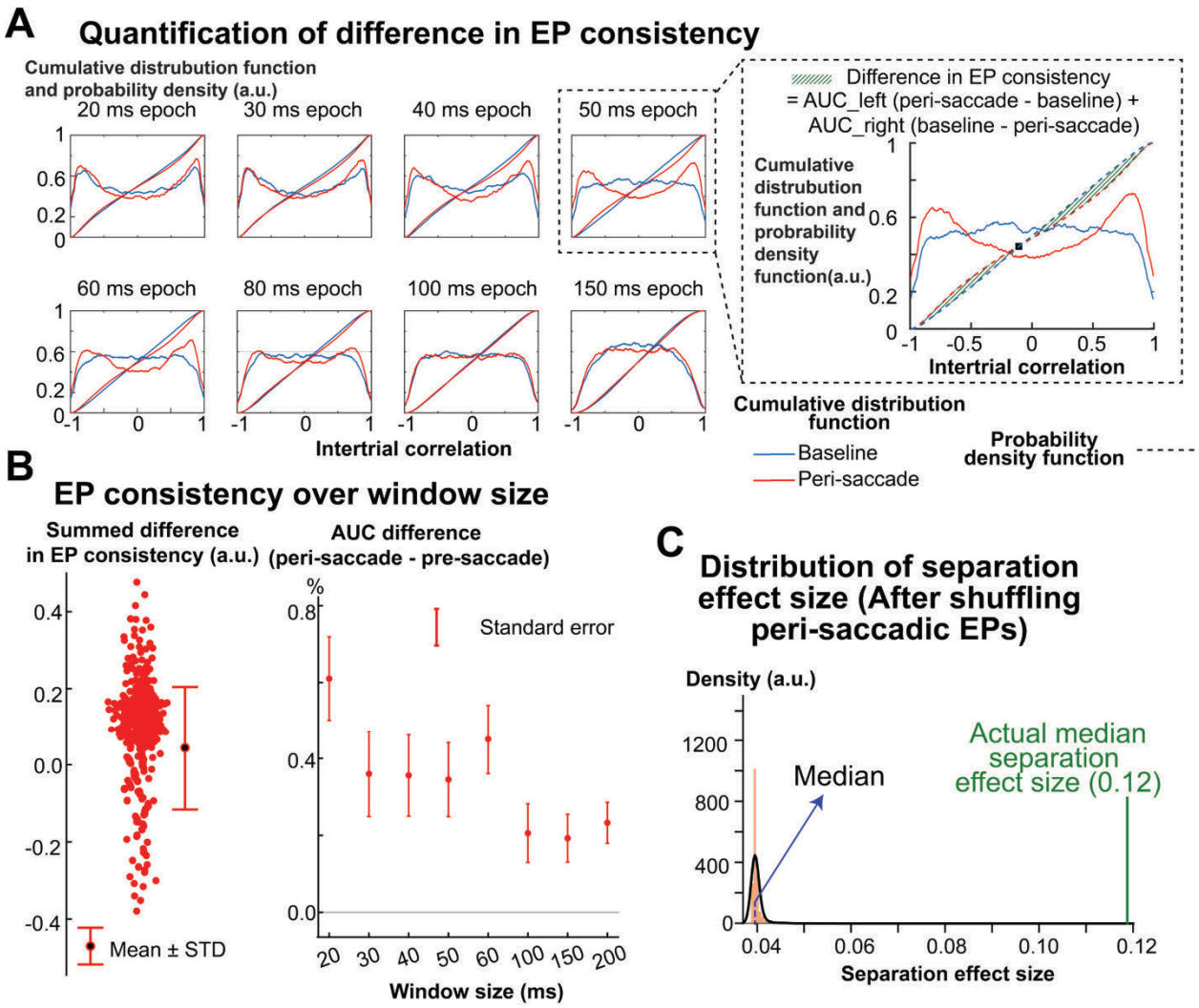


Figure 2 – Figure Supplement 4. Evoked potential consistency across time. To determine the time over which the EP waveform remains stable, we calculated the intertrial correlation distribution across a range of window sizes (20, 30, 40, 50, 60, 100, 150, 200 ms). We defined the pre-saccadic epoch starting from 1s prior to saccade offset, with the same length as the peri-saccadic epoch. **(A)** Cumulative distributions of intertrial correlation for an exemplary electrode (same electrode as shown in **Figure 2a**). As the window length increased, the degree of bimodality in the intertrial correlation distribution decreased. A bimodal distribution of intertrial correlations, with values closer to 1 or -1, indicates greater EP consistency. To quantify EP consistency differences between pre- and peri-saccadic epochs, we computed the difference in the area under the curve (AUC): 1. For correlations smaller than the crossing point (black rectangle), we calculated (peri-saccade AUC – pre-saccade AUC). 2. For correlations greater than the crossing point, we calculated (pre-saccade AUC – peri-saccade AUC). **(B)** EP consistency across window length. Similar to our main analysis (**Figure 2G**), we defined electrode contacts showing consistent EPs following saccades as those whose mean AUC of intertrial

correlation following saccades exceeded the mean pre-saccade AUC by two standard deviations. Across those electrode contacts, the summed difference in EP consistency was significantly greater than 0 ($t = 5.349$, $p < 0.001$, Cohen's $d = 0.274$, $n = 382$, two-sided t-test). Across all window sizes, the peri-saccadic EPs consistently exhibited greater waveform consistency compared to the pre-saccadic baseline. **(C)** Separation effect size following temporal shuffling of local field potentials. As a negative control, we disrupted the temporal structure by shuffling the local field potential data and repeated the analysis of the separation effect size, as defined in **Figure 2 – Figure Supplement 2a**. The separation effect size was the sum of the absolute AUC differences to the left and right of the crossing point between pre- and peri-saccadic cumulative distribution functions. The median separation between peri-saccadic and baseline EP was 0.04 in this negative control, significantly smaller than the observed separation effect size ($t = 65.086$, $p < 0.001$, Cohen's $d = 1.804$, two-sided t-test, $n = 2604$, compared to **Figure 2 – Figure Supplement 2B**), indicating our main findings remain robust.

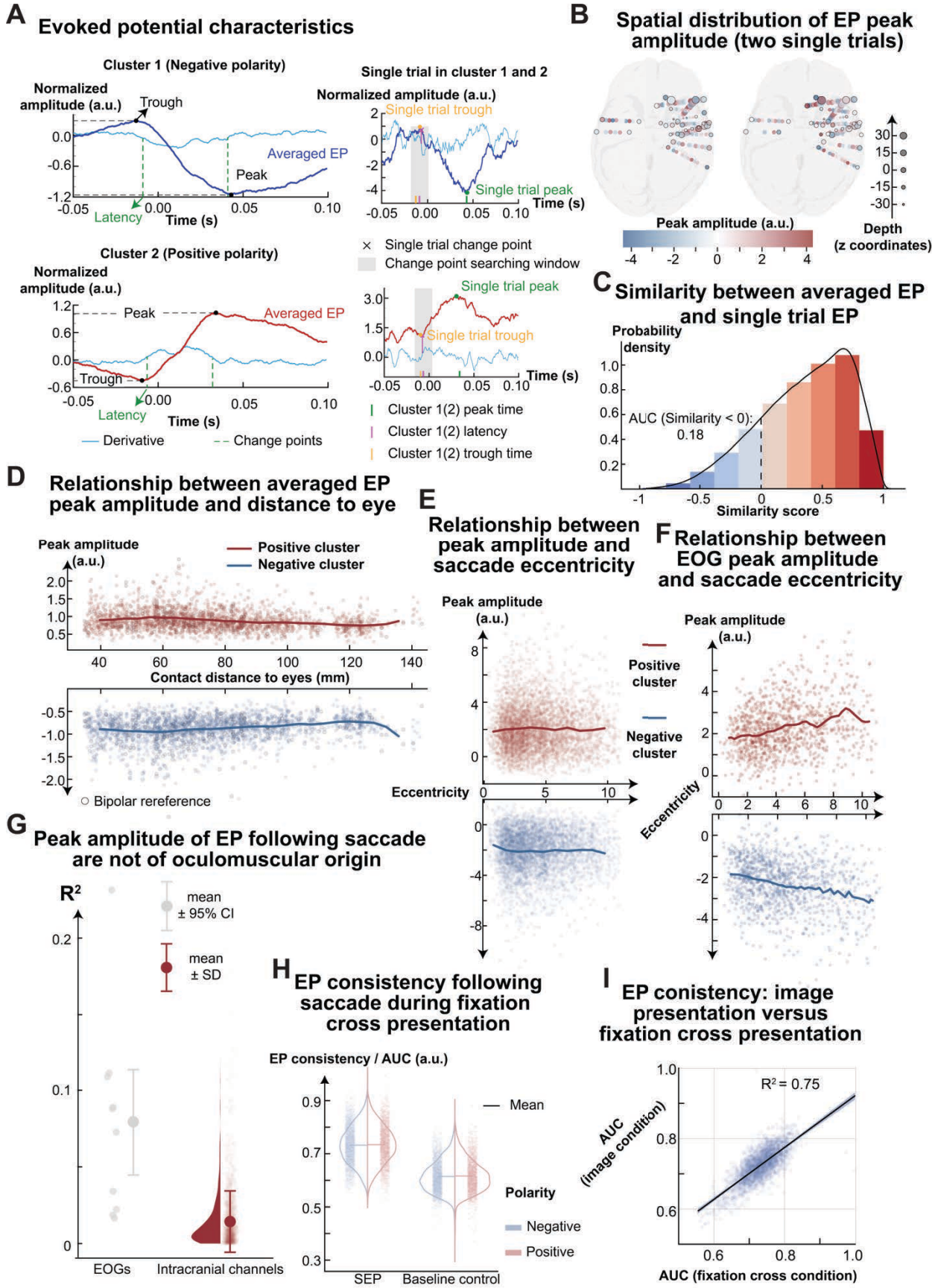

A
Evoked potential characteristics
Cluster 1 (Negative polarity)
Normalized amplitude (a.u.)
Trough
Averaged EP
Peak
Latency
Time (s)
Cluster 2 (Positive polarity)
Normalized amplitude (a.u.)
Peak
Averaged EP
Trough
Latency
Time (s)
Derivative
Change points
Single trial in cluster 1 and 2
Normalized amplitude (a.u.)
Single trial trough
Single trial peak
Time (s)
Single trial change point
Change point searching window
Single trial peak
Single trial trough
Time (s)
Cluster 1(2) peak time
Cluster 1(2) latency
Cluster 1(2) trough time
B
Spatial distribution of EP peak amplitude (two single trials)
Peak amplitude (a.u.)
Depth (z coordinates)
C
Similarity between averaged EP and single trial EP
Probability density
AUC (Similarity < 0): 0.18
Similarity score
D
Relationship between averaged EP peak amplitude and distance to eye
Peak amplitude (a.u.)
Positive cluster
Negative cluster
Contact distance to eyes (mm)
Bipolar rereference
E
Relationship between peak amplitude and saccade eccentricity
Peak amplitude (a.u.)
Eccentricity
F
Relationship between EOG peak amplitude and saccade eccentricity
Peak amplitude (a.u.)
Positive cluster
Negative cluster
Eccentricity
G
Peak amplitude of EP following saccade are not of oculomuscular origin
R²
mean ± 95% CI
mean ± SD
EOGs
Intracranial channels
H
EP consistency following saccade during fixation cross presentation
EP consistency / AUC (a.u.)
Mean
Polarity
Negative
Positive
SEP
Baseline control
I
EP conistency: image presentation versus fixation cross presentation
R² = 0.75
AUC (image condition)
AUC (fixation cross condition)

Figure 3. **SREPs are not of oculomuscular origin**. **(A)** Characterization of averaged and single-trial SREPs. We characterized EPs using latency, peak amplitude, trough amplitude, peak time, and trough time. After hierarchical clustering, we used MATLAB's *findchangepts* function to identify two change points in the derivative of EPs averaged across trials within a cluster. We defined latency as the first point and peak time as the time point within a 10-ms window around the second change point at which the z-scored SREP amplitude was maximal for positive clusters (or minimal for negative clusters). Similarly, we defined trough time as the time point within a 10-ms window around latency at which the z-scored SREP amplitude is minimal for positive clusters (or maximal for negative clusters). We identified latency for single-trial SREP as the change point around the latency identified from the corresponding averaged SREP. We identified the peak and trough amplitudes for single-trial SREP by finding the maximum and minimum SREP amplitudes around the corresponding cluster-level peak and trough times. **(B)** Spatial distribution of SREP amplitudes following individual saccades. This panel shows SREP peak amplitudes across intracranial electrode contacts for two saccades in a subject. Circle size indicates electrode depth, with smaller circles representing deeper electrode contacts. We did not observe any spatial pattern in polarity or peak amplitude. **(C)** Similarity between single-trial SREPs and the corresponding cluster-level averaged EPs, assessed using Spearman's correlation. Similarity distributions were skewed towards 1 (positive correlation), with an area under the curve of 0.812 for correlations greater than 0, compared to 0.181 for correlations less than 0, indicating that most single-trial EPs resemble their cluster-level averaged SREP. **(D)** Relationship between averaged EPs' peak amplitude and electrode-to-eye distance. The x-axis represents the distance from the recording electrode contact to the nearest eye. The line represents the average peak amplitude calculated within sliding windows of 10 mm, incremented by 5 mm steps. The distance explained 3.8% and 3.5% variance in peak amplitude of averaged SREPs for positive and negative clusters, respectively, across all electrode contacts. **(E)** Relationship between single-trial SREP amplitudes and saccade eccentricity. For both polarities, saccade eccentricity explained 0.0% variance in peak amplitude. Due to the large sample size, we show 2% of the data in the plot. **(F)** Relationship between single-trial oculomuscular signal peak amplitudes and saccade eccentricity. We derived oculomuscular signals (electrooculogram, EOG) extracted electrooculogram (EOG) signals from fronto-polar EEG scalp electrodes (Fp1 and Fp2) using bipolar re-referencing. Eccentricity explains 7.0% and 6.8% variance in oculomuscular signal peak amplitude for positive and negative clusters. **(G)** Comparison of correlation coefficients between saccade eccentricity and single-trial EP amplitudes for intracranial electrodes versus EOG. For each electrode contact, we calculated the correlation between eccentricity and single-trial EP amplitudes for this contact. The mean correlation coefficient was 0.014 for intracranial electrodes and 0.079 for oculomuscular signals. Welch's t-test for comparison

between unbalanced sample sizes showed that correlation coefficients for EOGs were significantly higher than intracranial local field potential signals (t = 3.154, p = 0.012, Cohen's d = 3.193, $n_{EOG} = 10$, $n_{EOG} = 2367$). For EOG correlation coefficients, the horizontal bars represent the bootstrapped 95% confidence interval (CI). **(H)** Consistency of SREP during fixation cross presentation, quantified using AUC as described in **Figure 2c**. During fixation cross presentation, SREP consistency was significantly higher than baseline (paired t-test, p<0.001, Cohen's d = 2.631, n = 2604). **(I)**. Correlation between SREP consistency during image presentation and during fixation cross presentation. Electrode contacts with consistent SREPs during image presentation also exhibited consistent SREPs following saccades during fixation cross presentation.

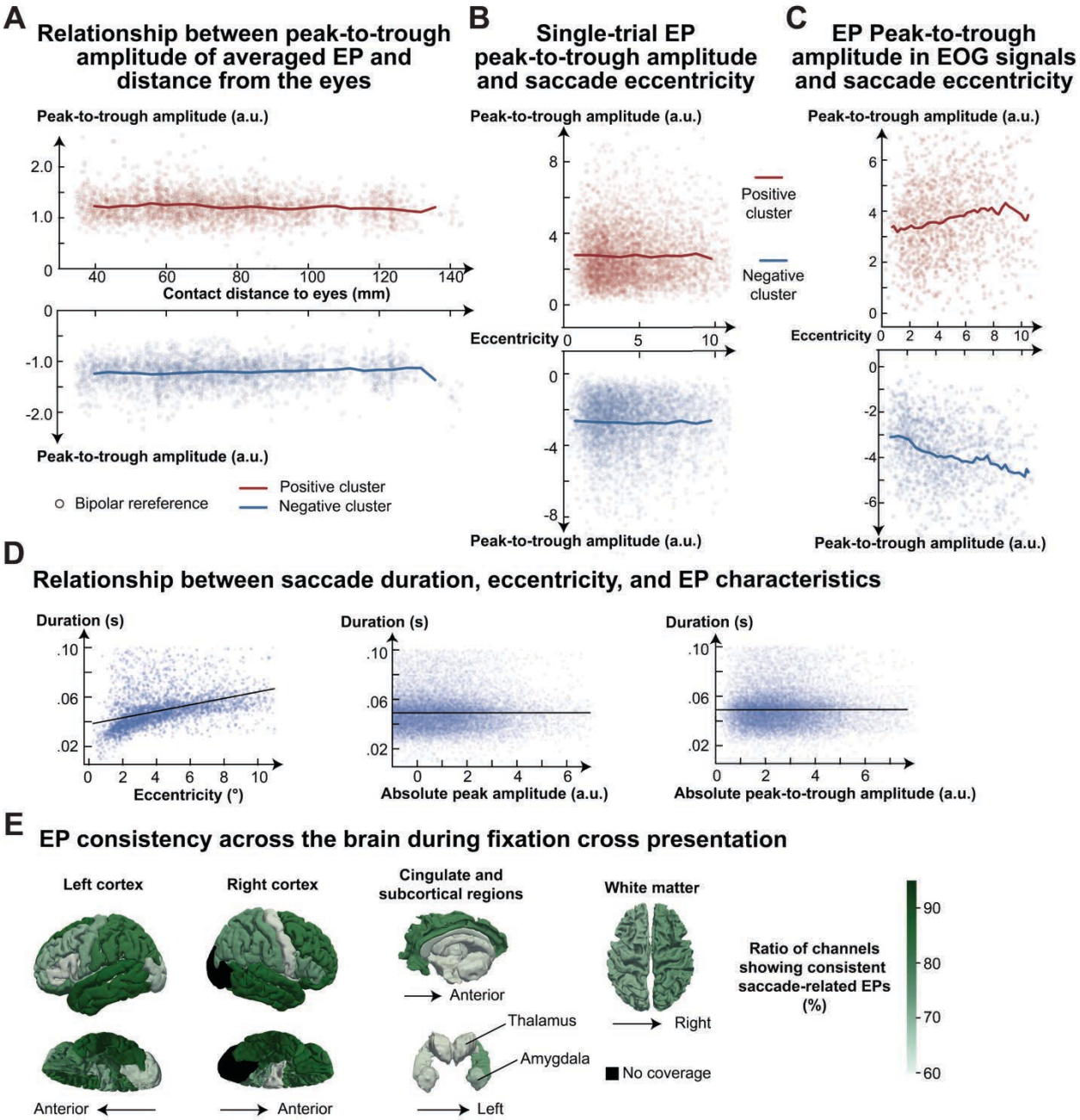


Figure 3 – Figure Supplement 1. SREPs are not oculomuscular signals. **(A)** Relationship between averaged SREPs' peak-to-trough amplitude and electrode-to-eye distance. The x-axis represents the distance from the recording electrode to the nearest eye. The line represents the average peak amplitude calculated within sliding windows of 10 mm, incremented by 5 mm steps. The distance explained 0.9% and 0.7% variance in peak to trough of averaged SREPs for positive and negative clusters, respectively, across all electrode contacts. **(B)** Relationship between single-trial SREP peak to trough and saccade eccentricity. For both polarities, saccade eccentricity explained 0.0% of the variance in peak-to-trough amplitude. **(C)** Relationship between single-trial

oculomuscular signal amplitudes and saccade eccentricity. We extracted electrooculogram (EOG) signals from fronto-polar EEG scalp electrodes (Fp1 and Fp2) using bipolar re-referencing. Eccentricity explains 2.7% and 8.2% variance in EP peak-to-trough amplitude for positive and negative clusters. **(D)** Relationship between saccade duration and eccentricity, SREP peak amplitude, and SREP peak to trough. Saccade duration and saccade eccentricity were positively correlated ($Pearson's\ r = 0.056$), while saccade duration was correlated less with SREP peak amplitude ($Pearson's\ r = -0.002$), or SREP peak to trough ($Pearson's\ r = 0.004$). **(E)** Response ratios during fixation cross presentation across the cerebrum. The response ratio is defined as the proportion of electrode contacts with AUC in the peri-saccadic condition above the mean + 2 standard deviation (SD) of baseline AUC. The AUC calculation is illustrated in **Figure 2C–D**. The brain was segmented into White Matter (Left, Right), Hippocampus (Left, Right), Amygdala (Left, Right), Basal Ganglia, Thalamus, Orbitofrontal (Left, Right), Inferior Frontal (Left, Right), Middle Frontal (Left, Right), Superior Frontal (Left, Right), Anterior Cingulate, Posterior Cingulate, Motor (Left, Right), Anterior Temporal (Left, Right), Middle Temporal (Left, Right), Superior Temporal (Left, Right), Inferior Temporal (Left, Right), Medial Temporal (Left, Right), Fusiform (Left, Right), Parietal (Left, Right), Occipital (Left, Right), and Insula (Left, Right).

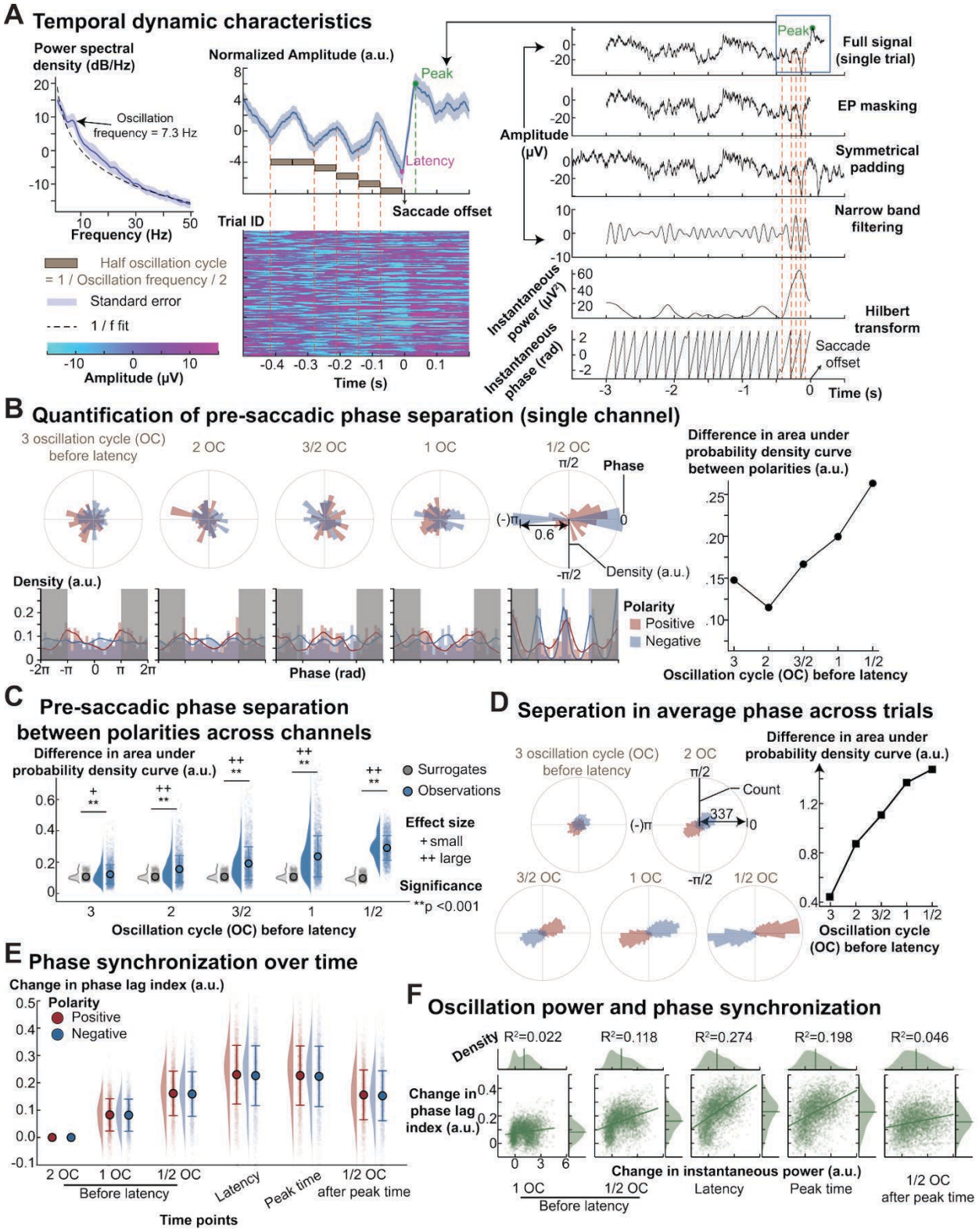


Figure 4. **Pre-saccadic oscillatory activity determines EP polarities following saccades**. **(A)** Identification of oscillatory activities. To account for the aperiodic activity, we fitted a one-over-frequency ($1/f$) curve. We defined the oscillation frequency as the

frequency at which the residual of the power spectral density from the $1/f$ fit was largest. The middle panel shows the averaged SREP and single-trial SREP time-locked to saccade offset. Based on the identified EP latency, we defined 5 time points of interest: 0.5, 1, 1.5, 2, and 3 oscillatory cycles (OC) before latency. The heatmap shows the voltage signals from single trials. We estimated the instantaneous phase and power using the Hilbert transform, applied to padded, narrowband-passed signal segments preceding EP latency (to avoid the potential influence of SREP in phase estimation). We only retained instantaneous power and phase corresponding to the masked, unpadded signal. **(B)** Pre-saccadic oscillation phase for the electrode contact shown in **A** and quantification of phase separation between SREPs of opposing polarities. The circular histograms show the distribution of phase for an electrode contact, colored by SREP polarity. To account for the circular nature of phase (i.e., -π is equal to π), we unwrapped phase values and extended the phase range from [$-\pi, \pi$] to [-2π, 2π]. We then estimated the probability density function and calculated the absolute difference in area under the probability density curve to quantify the phase separation between polarities. **(C)** Phase separation between polarities across electrode contacts. Each data point represents the difference in the area under the probability density function for one electrode contact. To assess whether the phases preceding SREP were different for positive and negative SREPs, we generated the surrogate (null) distribution by randomly permuting the polarity labels. The grey dots indicate the mean separation metric from the 50 permutations for each electrode contact. We used paired Wilcoxon signed-rank (two-sided, n = 1970 pairs per condition) to compare observed and null phase separation metrics across electrode contacts for each time point. All comparisons indicated statistically significant differences ($p < 0.001$). We quantified effect sizes using the rank biserial correlation, a non-parametric analog to Cohen's d that reflects the degree of separation between paired samples. Effect sizes increased progressively across conditions: from 0.204 in the earliest time point and 1.000 in the latest, suggesting strengthening of phase separation between polarities as time approached saccades. The rank-biserial correlation were 0.204, 0.581, 0.790, 0.901, 1.000 for the 5 time points. The data shown are from electrode contacts that exhibit both oscillatory activities and consistent EP following saccades. We defined contacts exhibiting oscillatory activities as those in which the largest deviation between the power spectral density and the fitted 1/f aperiodic curve occurred at a frequency below 30 Hz. This criterion excludes broadband activity (the elbow in the PSD) or high-frequency artifacts, such as 60 Hz line noise, resulting in 1970 electrode contacts (see **Figure 4 – figure supplement 1d**). **(D)** Average phase across trials for positive and negative SREPs. The circular histogram shows the distribution of phase for the two polarities averaged across trials for each electrode contact. The right panel shows increased separation in the average phase between polarities as the time approaches SREP latency. **(E)** Phase synchronization induced by SREPs. To quantify phase coherence around saccades, we estimated instantaneous power and phase using unmasked signals, in contrast to panel

A. The effect size (Cohen's d) between phase lag index at 2 OC before latency and subsequent time points (from 1 OC before latency to 1 OC after peak time) were 1.214, 1.625, 1.739, 1.704, and 1.427. The corresponding rank-based effect sizes (rank-biserial correlation) were 0.924, 0.980, 0.988, 0.982, 0.959. Phase lag index significantly increased after 2 OC before latency ($p < 0.001$, Wilcoxon paired test, $n = 5300$). **(F)** Relationship between instantaneous power and phase synchrony. Instantaneous power began increasing from 2 OC before the SREP latency and remained elevated (see **Figure 4 – figure supplement 1f**). The positive correlation between instantaneous power and phase synchrony was strongest at the SREP latency and weaker at other time points. The density plot at the top and right represents the unimodal distribution of instantaneous power and instantaneous phase, where the means of these distributions are marked.

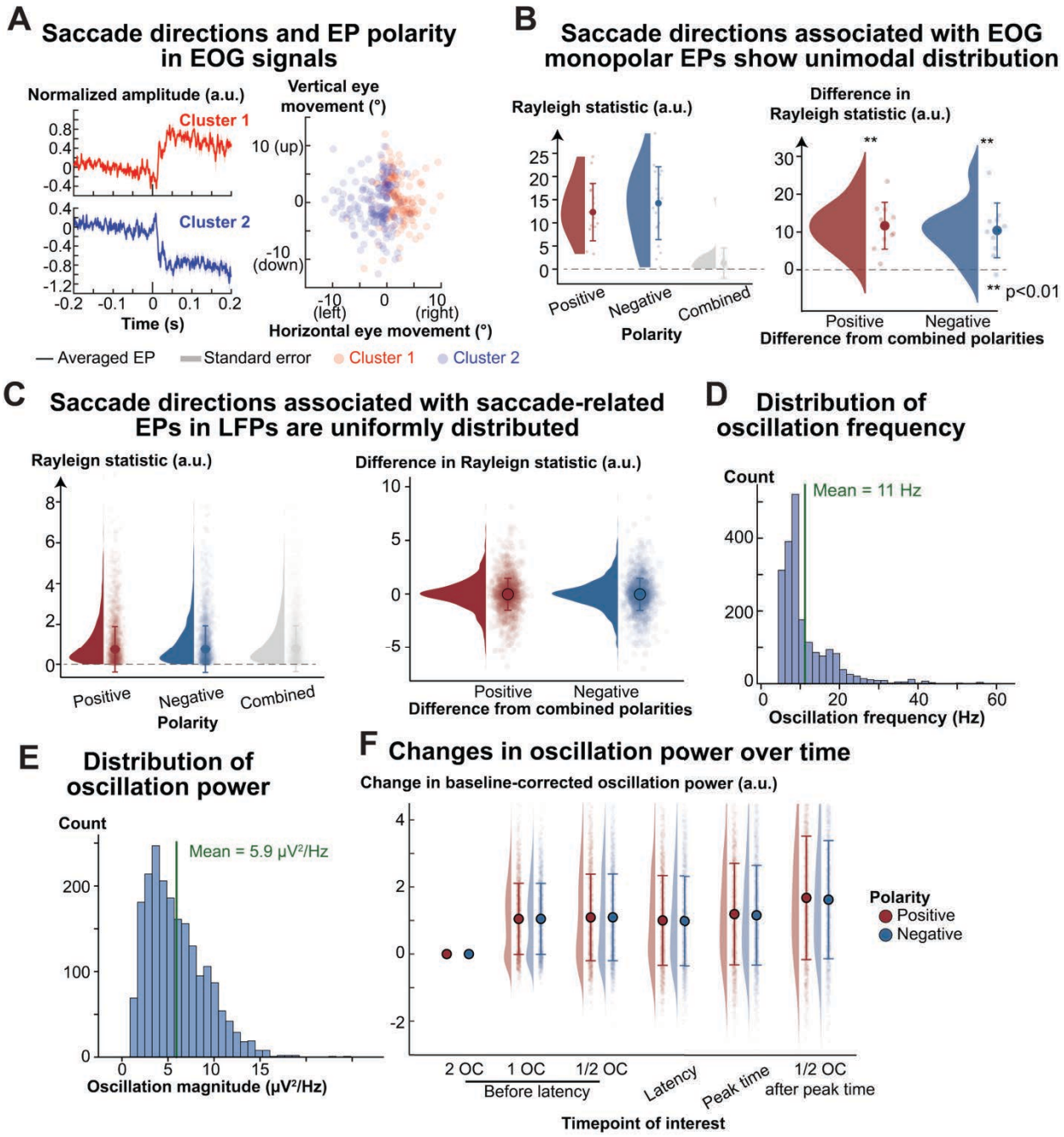


Figure 4 – Figure Supplement 1. Saccade direction is unimodally biased in monopolar EOG EPs but uniform in intracranial EPs. (A) Relationship between saccade direction and EP polarity in EOG signals. We applied the same clustering and EP characterization procedures to EOG signals. The left panel shows the positive and negative EOG EPs averaged across trials from a representative subject. Time 0 is the saccade offset. The right panel shows the saccade direction associated with the positive and negative EPs. The shaded area in the left panel represents the standard error. **(B)** Saccade directions were unimodally distributed in monopolar EOG EPs. Assuming EOG

signals captured oculomuscular artifacts during saccades, we expected that the circular distribution of saccade direction associated with monopolar EP has a unimodal peak. We used Rayleigh statistics to quantify this unimodal bias. Rayleigh statistics for monopolar EPs were higher than the Rayleigh statistics for combined EPs. Each data point in the plot represents Rayleigh statistics calculated from EOG signals in one subject during one task. We ensured the same sample size when calculating the Rayleigh statistics. We defined the resampling size as the 20th percentile of the number of trials (n = 32) for each subject, each task, and each electrode contact. We sampled the same number of saccade directions associated with monopolar EPs and saccade directions associated with the combined EPs. The right panel shows the distribution of Rayleigh statistic differences (monopolar – combined). For both polarities, differences in Rayleigh statistics were significantly greater than zero (For positive polarity, $t = 5.950$, $p < 0.001$, Cohen's d = 1.881, paired t-test, n = 15; for negative polarity, $t = 4.553$, $p = 0.001$, Cohen's d = 1.440, n = 18). In addition, the circular distribution of saccade direction was significantly different between positive and negative polarity ($p < 0.001$ for all subjects, 95% CI for W statistics = [33.223, 46.050], Mardia-Watson-Wheeler test). **(C)** Saccade directions associated with positive and negative SREPs in LFPs measured with stereoelectroencephalography were uniformly distributed. For both polarities, Rayleigh statistic differences (monopolar – combined) were centered at zero (For positive polarity, $t=-0.953$ $p = 0.341$, Cohen's d = -0.022, two-tailed t-test, n = 2068; for negative polarity, $t = -1.028$, $p = 0.304$, Cohen's d = -0.023, n = 2058). To estimate the difference between the saccade directions associated with positive and negative SREPs, we sampled the same number of saccade directions associated with monopolar EPs. Consistently, we defined the resampling size as the 20th percentile of the number of trials (n = 62) for each subject, each task, and each electrode contact. Across subjects, tasks, and electrode contacts, we found no difference in saccade direction between polarities (95% CI for W statistics = [1.926, 2.107], the corresponding 95% CI for p-value is [0.486, 0.512], Mardia-Watson-Wheeler test). **(D)** Peak oscillation frequencies in intracranial signals. The mean peak oscillatory frequency is 11 Hz. To minimize potential confounds of line noise, we identified peak oscillation frequency in the 1 – 58 Hz range from the difference between the power spectral density and the one-over-frequency fit. **(E)** Oscillation power in intracranial signals. Measured as the difference in power spectral density from a one-over-frequency fit at the oscillatory frequency. **(F)** Changes in oscillation power (i.e., instantaneous power at the oscillatory frequency) over time. To enable cross-subject and cross-contact comparisons, we normalize oscillation power to power at two oscillation cycles before the EP latency. The effect size (Cohen's d) between phase lag index at 2 OCs before latency and subsequent time points (from 1 OC before latency to 1 OC after peak time) were 0.894, 0.780, 0.694, 0.742, and 0.860 (n = 5300).

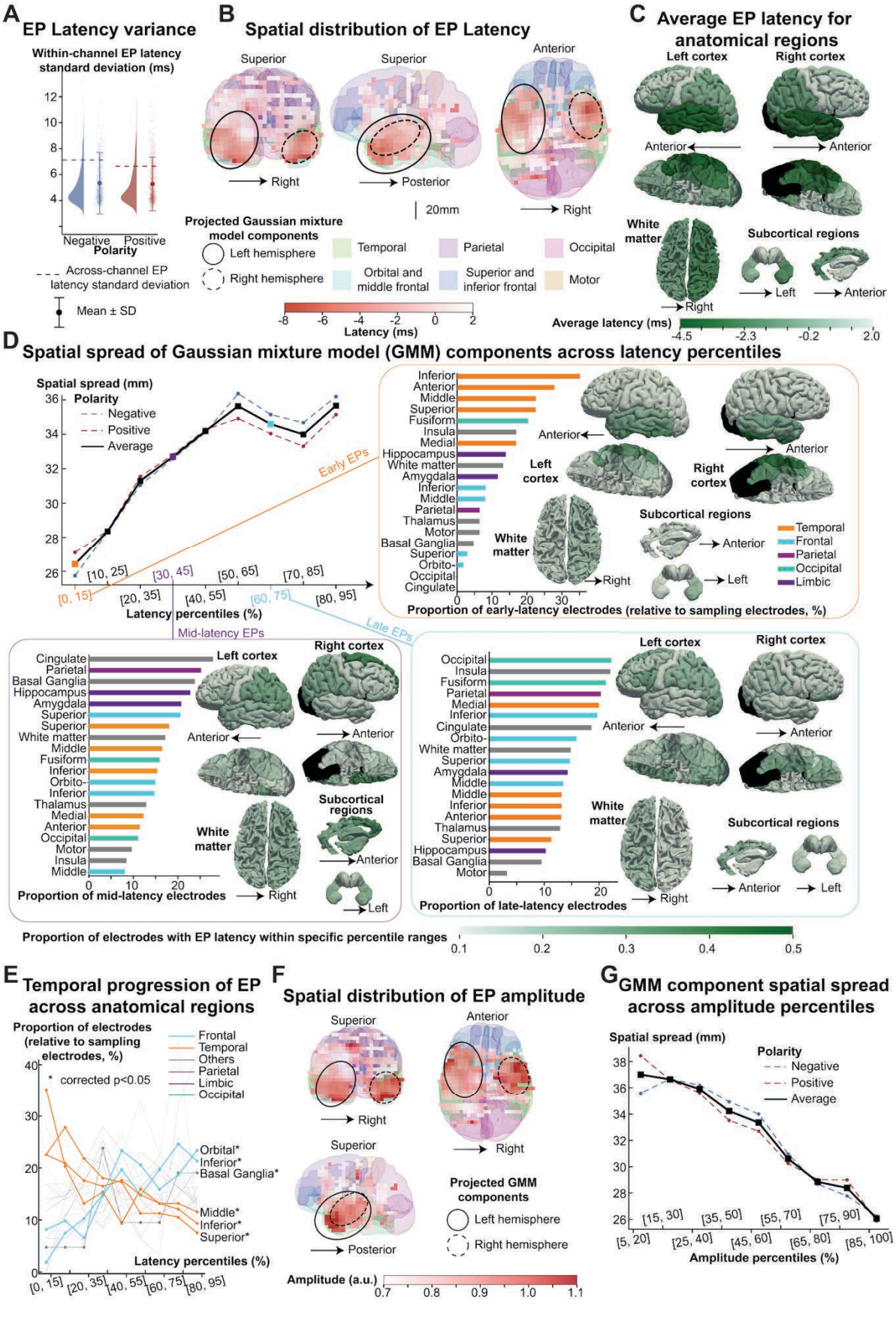

A
EP Latency variance
Within-channel EP latency standard deviation (ms)
Negative
Positive
Polarity
Across-channel EP latency standard deviation
Mean ± SD
B
Spatial distribution of EP Latency
Superior
Superior
Anterior
Right
Posterior
20mm
Right
Projected Gaussian mixture model components
Left hemisphere
Right hemisphere
Temporal
Parietal
Occipital
Orbital and middle frontal
Superior and inferior frontal
Motor
Latency (ms)
C
Average EP latency for anatomical regions
Left cortex
Right cortex
Anterior
Anterior
White matter
Subcortical regions
Left
Anterior
Right
Average latency (ms)
D
Spatial spread of Gaussian mixture model (GMM) components across latency percentiles
Spatial spread (mm)
Polarity
Negative
Positive
Average
Early EPs
Mid-latency EPs
Late EPs
Latency percentiles (%)
Proportion of early-latency electrodes (relative to sampling electrodes, %)
Proportion of mid-latency electrodes
Proportion of late-latency electrodes
Temporal
Frontal
Parietal
Occipital
Limbic
Proportion of electrodes with EP latency within specific percentile ranges
E
Temporal progression of EP across anatomical regions
Proportion of electrodes (relative to sampling electrodes, %)
* corrected p<0.05
Frontal
Temporal
Others
Parietal
Limbic
Occipital
Orbital*
Inferior*
Basal Ganglia*
Middle*
Inferior*
Superior*
Latency percentiles (%)
F
Spatial distribution of EP amplitude
Superior
Anterior
Right
Right
Superior
Posterior
Projected GMM components
Left hemisphere
Right hemisphere
Amplitude (a.u.)
G
GMM component spatial spread across amplitude percentiles
Spatial spread (mm)
Polarity
Negative
Positive
Average
Amplitude percentiles (%)

Figure 5. **Spatiotemporal dynamics of saccade-related EPs**. **(A)** Comparison between intra-contact SREP latency variance and inter-contact SREP latency variance. The averaged standard deviation of intra-contact SREP latencies was 5.3 ms for negative and positive EP. The corresponding variance of intra- contact SREP latencies was 34 μs^2 and 32 μs^2 for negative and positive EP, respectively. The standard deviation of intra- contact EP latencies was significantly lower compared to inter- contact EP latency of 7.1 ms and 6.6 ms for negative and positive EPs (i.e., variance of 51 μs^2 and 44 μs^2) (N = 786, W statistics = 24141, $p < 0.001$, Rank-biserial correlation= -0.766 for negative polarity; N = 797, W statistics = 34836, $p < 0.001$, Rank-biserial correlation= -0.669 for positive polarity; Wilcoxon signed-rank test). To estimate intra-contact SREP latency variance, we included single trials whose SREP waveforms were similar to the averaged SREP waveform, defined as having a correlation coefficient above the median. We resampled the single-trial EP latency for each contact to the 50th percentile of the trial count for all electrode contacts. To estimate inter-contact SREP latency variance, we resampled electrode contacts to have equal sample sizes, matching the median of per-contact trial counts. We repeated the process 100 times and plotted the mean inter-contact SREP latency variance using dashed lines. **(B)** Spatial distribution of SREP latencies across electrode contacts. Similar to **Figure 2 – Figure supplement 3**, we projected the 3D brain surface and electrode coordinates onto a 2D plane. We overlay a heatmap to represent the spatial pattern of EP latencies. To generate the heatmap, we discretized this 2D space into a regular grid with a 5 mm resolution. We computed the local average SREP latency within a 10 mm radius for each grid element. Heatmap elements that have no nearby sampling electrodes are fully transparent. To analyze the spatial dynamics of SREP over time, we stratified electrodes by SREP latency percentiles. For each percentile-defined subset, we applied a 3D Gaussian Mixture Model (GMM) to identify two spatial clusters of electrode contacts. We extracted the cluster centroids in the left and right hemispheres. We quantified the spatial spread based on standard deviations along the x, y, and z axes. The solid and dashed ellipses represent the two GMM components identified from electrode contacts exhibiting EP latency in the 0 – 15th percentile range. **(C)** Average SREP latency across anatomical regions. We parcellate the brain into White Matter (Left, Right), Hippocampus (Left, Right), Amygdala (Left, Right), Basal Ganglia, Thalamus, Orbitofrontal (Left, Right), Inferior Frontal (Left, Right), Middle Frontal (Left, Right), Superior Frontal (Left, Right), Anterior Cingulate, Posterior Cingulate, Motor (Left, Right), Anterior Temporal (Left, Right), Middle Temporal (Left, Right), Superior Temporal (Left, Right), Inferior Temporal (Left, Right), Medial Temporal (Left, Right), Fusiform (Left, Right), Parietal (Left, Right), Occipital (Left, Right), and Insula (Left, Right). Colors denote the mean EP latency, averaged across polarities and electrode contacts localized to each anatomical region. **(D)** Relationship between spatial spread and SREP latency percentile ranges. We quantified spatial spread as the square root of the sum of variances along the X, Y, and Z axes (i.e., the Euclidean norm of the standard deviations). For illustration, we

selected early, mid, and late percentile ranges ([0–15], [30–45], and [60–75]) and plotted the proportion of electrodes showing EP with latency within each range. We calculated the proportion as the ratio between the number of electrodes in a region with latency within a percentile range and the total number of electrodes localized in that region. The accompanying 3D brain plots show the anatomical distribution of these proportions across regions. **(E)** Temporal progression of EP across anatomical regions. Each data point represents the proportion of electrodes for an anatomical region that showed EP at a specific latency percentile range. Thick lines represent regions where the proportion changed monotonically across latency percentiles based on Spearman's correlation. We found that orbital ($\rho = 0.82$, Bonferroni corrected asympototic p = 0.136, number of comparisons = 20) and inferior ($\rho = 0.89$, $corrcted\ p = 0.025$) frontal regions had a lower proportion of early post-saccadic EPs and a higher proportion of late EPs. Middle ($\rho = -0.97$, $corrected\ p < 0.001$), inferior ($\rho = -0.74, corrcted\ p = 0.455$), and superior ($\rho = -0.98, corrcted\ p < 0.001$) temporal regions showed the opposite pattern. **(F)** Spatial distribution of EP peak amplitude across electrodes. **(G)** Relationship between spatial spread and EP peak amplitude percentile ranges.

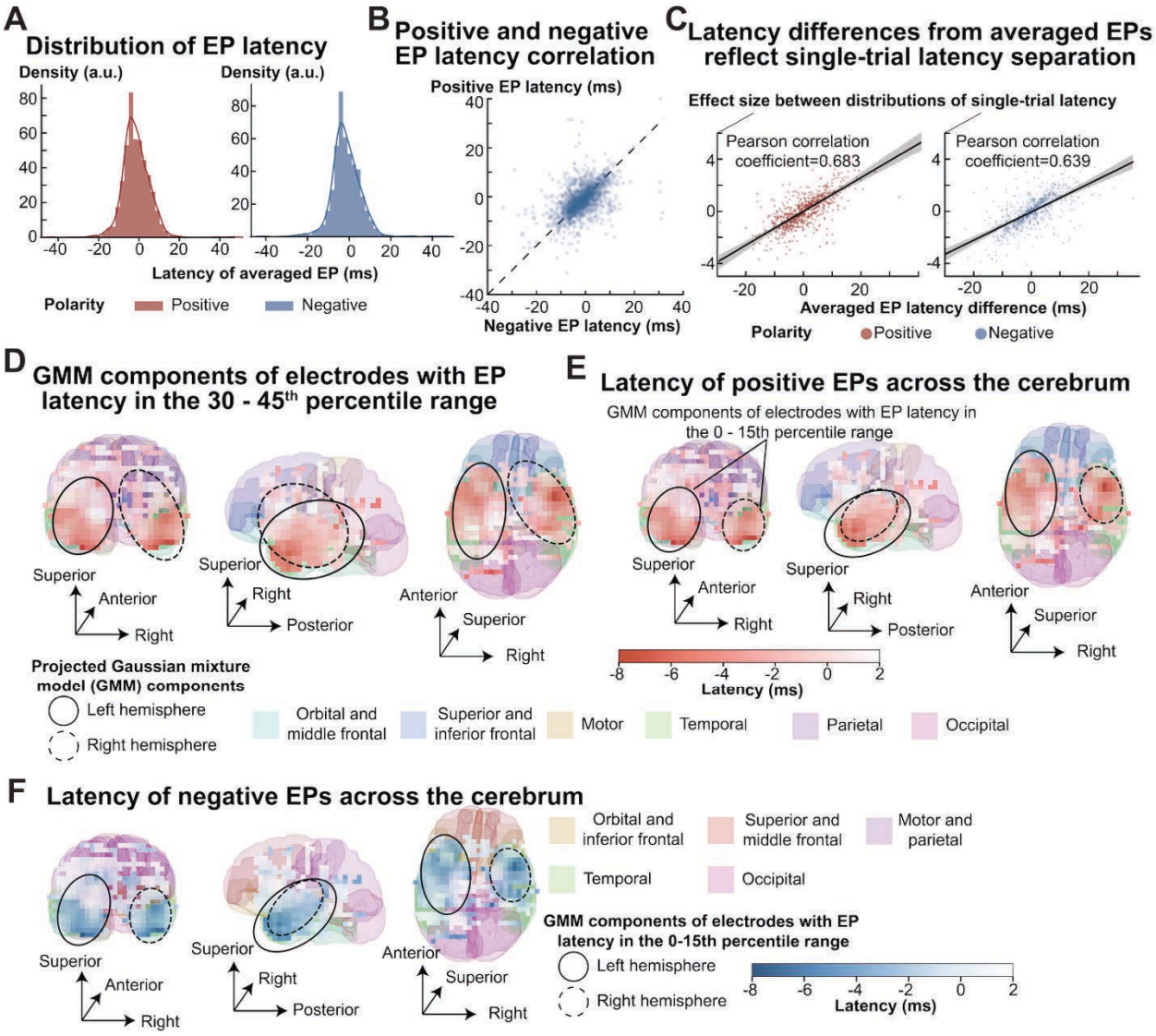


Figure 5 – Figure supplement 1. **SREP latency across the cerebrum**. **(A)** Averaged SREP latency. The mean latency was -1 ms for both positive and negative polarities. The standard deviation was 7.5 ms for positive SREPs and 7.7 ms for negative SREPs, respectively. **(B)** Correlation between latencies of averaged EPs with positive and negative polarities. They were positively correlated (Pearson's correlation coefficient = 0.436, $p < 0.001$, exact test for Pearson's correlation, $n = 1558$). **(C)** Latency differences between trial-averaged SREPs across electrode contact pairs positively correlated with the effect size of latency differences estimated from single-trial EPs ($p < 0.001$, exact test for Pearson's correlation, i.e., t-test). Given the large number of possible electrode contact pairs, we randomly sampled 2,000 contact pairs per polarity to reduce computational cost. We randomly sampled 2000 pairs of electrode contacts for each polarity. Illustration of two Gaussian Mixture Model (GMM) components estimated from electrodes with EP latency in the 30 – 45th percentile range. Each ellipse represents a GMM component,

derived from the centroid and standard deviations along the three spatial axes of the 3D Gaussian distributions. The heatmap represents the average EP latency averaged across polarities. The underlying projected 3D brain model provides an anatomical reference. As in **Figure 2 – Figure supplement 3**, the color of each heatmap element (5 mm by 5 mm) represents the average latency across electrode contacts localized within 10 mm of that element. **(E)** The spatial distribution of positive EP latency. **(F)** The spatial distribution of negative EP latency. Darker blue in the heatmap denotes earlier SREPs. The color of the underlying 3D brain is adapted from **E**, to avoid overlap with the blue-to-white latency heatmap. In **E** and **F,** the solid and dashed ellipses represent left and right GMM components estimated from the spatial location of electrodes with EP latency in the 0 – 15th percentile range.

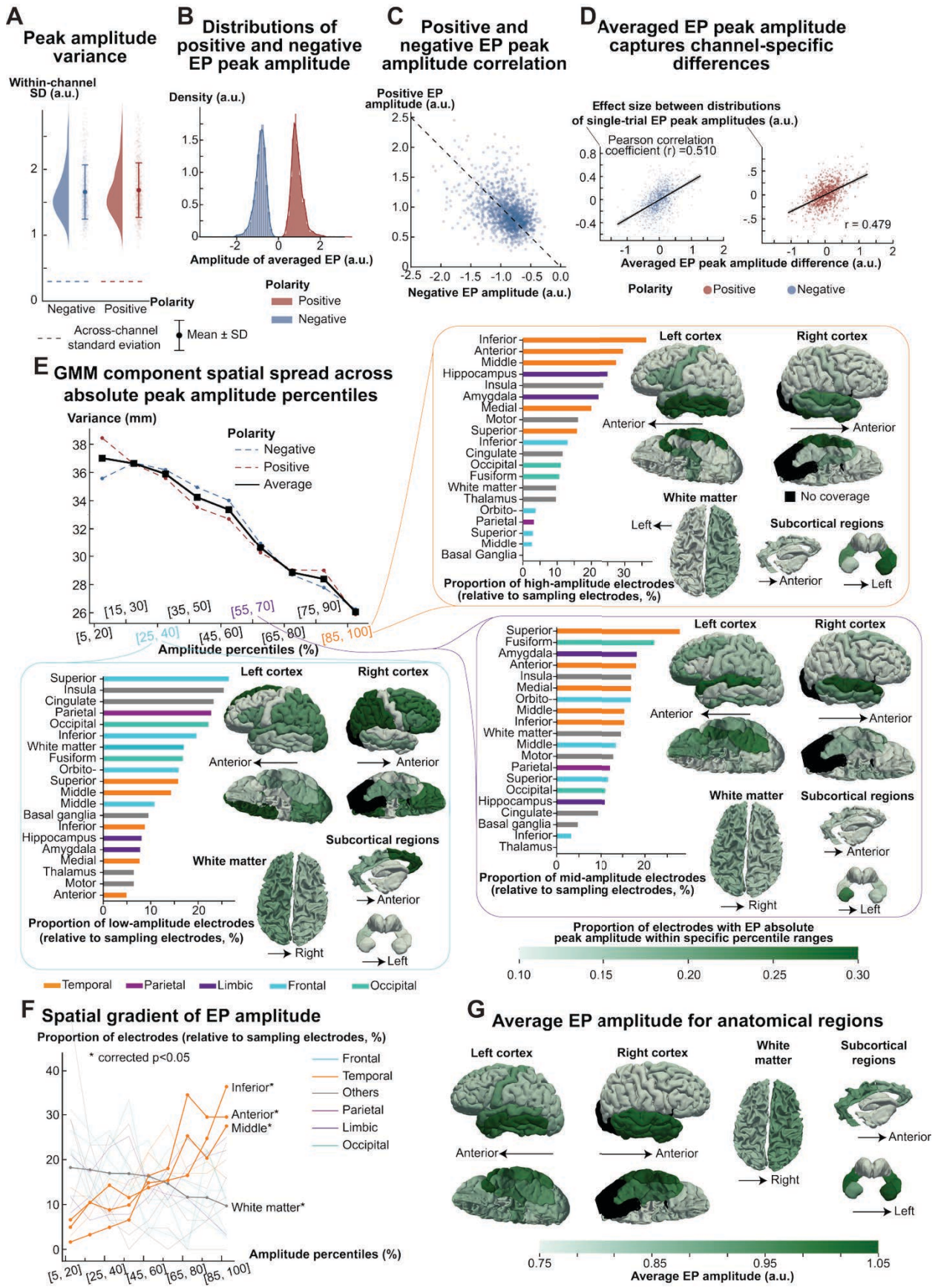

A
Peak amplitude variance
Within-channel SD (a.u.)
Negative
Positive
Polarity
Across-channel standard eviation
Mean ± SD
B
Distributions of positive and negative EP peak amplitude
Density (a.u.)
Amplitude of averaged EP (a.u.)
Polarity
Positive
Negative
C
Positive and negative EP peak amplitude correlation
Positive EP amplitude (a.u.)
Negative EP amplitude (a.u.)
D
Averaged EP peak amplitude captures channel-specific differences
Effect size between distributions of single-trial EP peak amplitudes (a.u.)
Pearson correlation coefficient (r) =0.510
r = 0.479
Averaged EP peak amplitude difference (a.u.)
Polarity
Positive
Negative
E
GMM component spatial spread across absolute peak amplitude percentiles
Variance (mm)
Polarity
Negative
Positive
Average
[5, 20]
[15, 30]
[25, 40]
[35, 50]
[45, 60]
[55, 70]
[65, 80]
[75, 90]
[85, 100]
Amplitude percentiles (%)
Left cortex
Right cortex
Anterior
White matter
Left
Subcortical regions
No coverage
Proportion of high-amplitude electrodes (relative to sampling electrodes, %)
Proportion of low-amplitude electrodes (relative to sampling electrodes, %)
Proportion of mid-amplitude electrodes (relative to sampling electrodes, %)
Right
Temporal
Parietal
Limbic
Frontal
Occipital
Proportion of electrodes with EP absolute peak amplitude within specific percentile ranges
F
Spatial gradient of EP amplitude
Proportion of electrodes (relative to sampling electrodes, %)
* corrected p<0.05
Frontal
Temporal
Others
Parietal
Limbic
Occipital
Inferior*
Anterior*
Middle*
White matter*
Amplitude percentiles (%)
G
Average EP amplitude for anatomical regions
Left cortex
Right cortex
White matter
Subcortical regions
Anterior
Right
Left
Average EP amplitude (a.u.)

Figure 5 – Figure supplement 2. **Spatial distribution of SREP peak amplitude**. **(A)** Comparison between intra-contact EP peak amplitude variance and inter- contact SREP amplitude variance. The average standard deviation of intra- contact SREP peak amplitude was 1.656 and 1.685 for negative and positive SREP, respectively. The standard deviation of intra-contact SREP amplitudes was significantly higher compared to inter-contact SREP latency of 0.299 and 0.298 (N = 786, W statistics = 0, $p < 0.001$, Rank-biserial correlation = 1 for negative polarity; N = 797, W statistics = 0, $p < 0.001$, Rank-biserial correlation = 1 for positive polarity; Wilcoxon signed-rank test). **(B)** Averaged SREP peak amplitudes. The mean peak amplitude was 0.897 and -0.883 for positive and negative polarities, respectively. The standard deviation was 0.296 for positive and negative SREPs. **(C)** Correlation between positive and negative averaged SREP peak amplitudes. They were negatively correlated (Pearson's correlation coefficient = -0.365, $p < 0.001$, exact test for Pearson's correlation). **(D)** Correlation between peak amplitude differences between trial-averaged SREPs across electrode contact pairs and the effect size of peak amplitude differences estimated from single-trial SREPs. They were positively correlated ($p < 0.001$, exact test for Pearson's correlation). **(E)** Relationship between spatial spread and absolute EP peak amplitude percentile ranges (related to **Figure 5D**). For illustration, we selected high, mid, and low amplitude percentile ranges ([0–15], [30–45], and [60–75]) and plotted the proportion of electrodes showing EP with amplitude within each range. We calculated the proportion as the ratio between the number of electrodes in a region with amplitude within a percentile range and the total number of electrodes localized in that region. The accompanying 3D brain plots show the anatomical distribution of these proportions across regions. **(F)** Proportion of electrodes with absolute SREP amplitude at a specific percentile range per region across absolute amplitude percentile bins. Thick lines represent regions where the proportion changed monotonically across amplitude percentile ranges based on Spearman's correlation. We found that anterior ($\rho = 0.95$, Bonferroni corrected asympototic $p = 0.002$, number of comparisons = 20), inferior ($\rho = 0.95$, $corrcted\ p = 0.001$),and middle ($\rho = 0.93$, $corrcted\ p = 0.005$) temporal regions had a lower proportion of electrodes showing post-saccadic EPs with low absolute amplitude and a higher proportion of electrodes showing post-saccadic EPs with high absolute amplitude. White matter ($\rho = -1$, $corrcted\ p < 0.001$) showed the opposite pattern. **(G)** Spatial distribution of absolute SREP peak amplitude averaged across electrode contacts localized within an anatomical region.

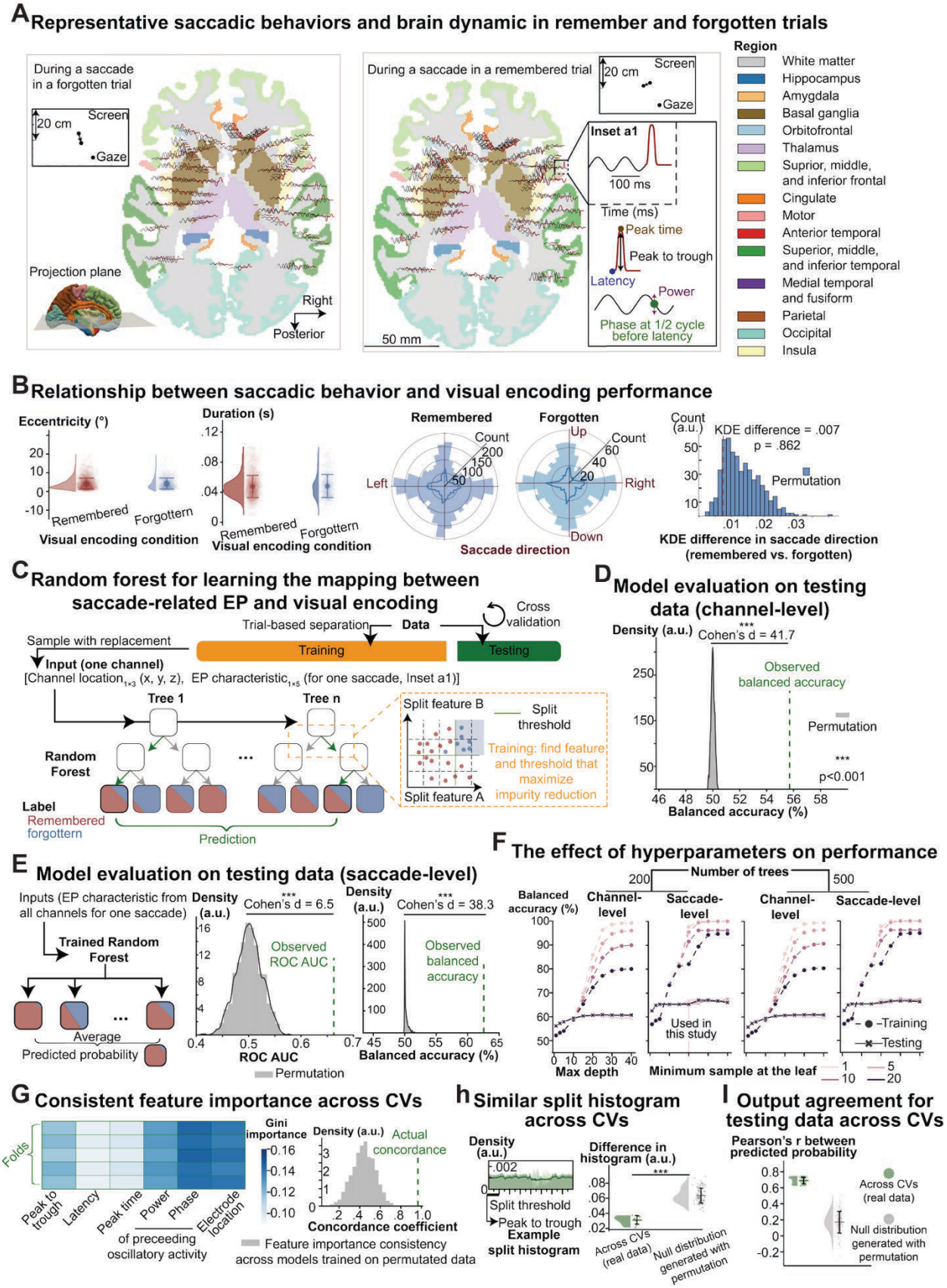

A Representative saccadic behaviors and brain dynamic in remember and forgotten trials
During a saccade in a forgotten trial
20 cm
Screen
Gaze
Projection plane
Right
Posterior
During a saccade in a remembered trial
Inset a1
100 ms
Time (ms)
Peak time
Peak to trough
Latency
Power
Phase at 1/2 cycle before latency
50 mm
Region
White matter
Hippocampus
Amygdala
Basal ganglia
Orbitofrontal
Thalamus
Suprior, middle, and inferior frontal
Cingulate
Motor
Anterior temporal
Superior, middle, and inferior temporal
Medial temporal and fusiform
Parietal
Occipital
Insula
B Relationship between saccadic behavior and visual encoding performance
Eccentricity (°)
Remembered
Forgottern
Visual encoding condition
Duration (s)
Remembered
Forgottern
Visual encoding condition
Remembered
Forgotten
Count
Left
Right
Up
Down
Saccade direction
Count (a.u.)
KDE difference = .007
p = .862
Permutation
KDE difference in saccade direction (remembered vs. forgotten)
C Random forest for learning the mapping between saccade-related EP and visual encoding
Cross validation
Trial-based separation
Data
Training
Testing
Sample with replacement
Input (one channel)
[Channel location1×3 (x, y, z), EP characteristic1×5 (for one saccade, Inset a1)]
Tree 1
Tree n
Random Forest
Label
Remembered
forgottern
Prediction
Split feature B
Split threshold
Training: find feature and threshold that maximize impurity reduction
Split feature A
D Model evaluation on testing data (channel-level)
Density (a.u.)
Cohen's d = 41.7
Observed balanced accuracy
Permutation
p<0.001
Balanced accuracy (%)
E Model evaluation on testing data (saccade-level)
Inputs (EP characteristic from all channels for one saccade)
Trained Random Forest
Average
Predicted probability
Density (a.u.)
Cohen's d = 6.5
Observed ROC AUC
ROC AUC
Permutation
Density (a.u.)
Cohen's d = 38.3
Observed balanced accuracy
Balanced accuracy (%)
F The effect of hyperparameters on performance
Number of trees
200
500
Balanced accuracy (%)
Channel-level
Saccade-level
Channel-level
Saccade-level
Used in this study
Training
Testing
Max depth
Minimum sample at the leaf
1
5
10
20
G Consistent feature importance across CVs
Folds
Peak to trough
Latency
Peak time
Power
Phase
Electrode location
of preceeding oscillatory activity
Gini importance
Density (a.u.)
Actual concordance
Concordance coefficient
Feature importance consistency across models trained on permutated data
h Similar split histogram across CVs
Density (a.u.)
Split threshold
Peak to trough
Example split histogram
Difference in histogram (a.u.)
Across CVs (real data)
Null distribution generated with permutation
I Output agreement for testing data across CVs
Pearson's r between predicted probability
Across CVs (real data)
Null distribution generated with permutation

Figure 6. **Machine learning reveals predictive mappings between saccade-related EPs and visual encoding**. **(A)** Illustration of EP dynamics during a remembered and a forgotten trial. Each trace plotted on top of a horizontal brain slice represents the EP waveform and the preceding oscillatory activity recorded in a electrode contact. The color of the brain slice denotes anatomical regions. We reconstructed EP waveforms as sigmoidal rise and fall based on latency, peak time, and peak-to-trough identified in the specific trials. The sigmoidal rise starts at the latency and peaks at the peak time. We reconstruct the preceding oscillatory activity using a 10 Hz sine wave parameterized by instantaneous power and phase at half a cycle before latency. We later trained linear and nonlinear models to predict visual encoding using five features characterizing the EP waveform and preceding oscillatory activity at each electrode contact. The two insets at the top show fixation before and after a saccade, with arrows indicating saccade direction. **(B)** Relationship between saccadic behavior and visual encoding performance. From left to right, we show the distribution of saccade eccentricity, saccade duration, and saccade direction in remembered and forgotten trials. The rightmost panel shows the difference in circular distribution of saccade direction between remembered and forgotten trials. To quantify the difference in saccade direction, we estimated circular kernel density functions for remembered and forgotten trials and computed the absolute area under the difference between the two density curves. The red dotted line represents the actual difference in saccade direction between remembered and forgotten trials. The blue histogram represents the null distribution of saccade direction differences from permuted trial conditions (remembered vs. forgotten). **(C)** Model architecture for learning the mapping from saccade-related EP to visual encoding. The input to the random forest is a vector including the spatial location of electrode contacts and SRND parameters (e.g., SREP latency, peak-to-trough, peak time, phase and power of preceding oscillatory activity). As the remembered and forgotten trials are imbalanced, we resampled both classes to achieve a balanced input. In addition, we added normally distributed spatial jitters to electrode contact location to enable learning mapping specific to spatial location while preventing associating the exact electrode contact location with the trial label. For prediction, each decision tree outputs a probability for the remembered and forgotten classes that reflects the proportion of training samples from each class reaching the corresponding leaf. The final predicted probability is the average of these vectors across all trees. **(D)** Balanced accuracy of the random forest in predicting visual encoding performance on testing data. We performed 200 permutations to generate a null distribution of balanced accuracy. The observed balanced accuracy was 55.7%, exceeding all values in the null distribution. **(E)** Saccade-level model performance. Using these saccade-level predicted probabilities, we computed the area under the receiver operating characteristic curve (ROC AUC). An illustration of calculating ROC AUC is shown in **Figure 2C**. The observed ROC AUC on test data (0.662) exceeded all values from a null distribution

generated by training and evaluating models on data with permuted labels. The rightmost panel shows the observed saccade-level balanced accuracy (62.6%), compared to the null distribution. To calculate the saccade-level balanced accuracy, we determined the optimal threshold, sweeping from 0 to 1, that maximized balanced accuracy. We applied this process to both observed data and data with permutated labels. **(F)** Effect of hyperparameters on model performance during training and testing. We evaluated how different hyperparameter settings affected the performance of random forests. The hyperparameter includes the number of trees, the maximum tree depth, and the minimum sample size at the leaf. In general, deeper trees with a small sample size can learn more complex patterns but risk overfitting. In each panel, color represents the minimum sample at the leaf, the x-axis represents the maximum depth, and the y-axis represents balanced accuracy. From left to right, the four panels show: (1) electrode contact-level performance using 200 trees, (2) saccade-level performance using 200 trees, (3) contact-level performance using 500 trees, and (4) saccade-level performance using 500 trees. In this study, we used a default hyperparameter setting of 200 decision trees, a maximum depth of 20, and a minimum leaf size of 1. This set of hyperparameters achieved moderate balanced accuracy during training and moderate generalization on test data, without evidence of overfitting. Therefore, we used the default parameters for consistency and computational efficiency, although performance on test data slightly improved with 500 decision trees. **(G)** Similar mapping from SRND parameters to visual encoding outcome across cross-validations (CVs). The left panel shows the feature importance based on the Gini impurity reduction across the 5-fold validations. We used Kendall's coefficient of concordance to quantify the concordance of feature importance across CVs and across permutated data. The right panel shows that the concordance coefficient across CVs (Kendall's W = 0.966) for the actual data exceeds all values in the null distribution (mean Kendall's W = 0.439, SD = 0.110). **(H)** Consistent decision boundaries across cross-validation folds. For each model trained on the actual data, we computed the histogram of split thresholds for every feature. To assess consistency, we compared split histograms between pairs of folds and quantified their similarity using the area under the absolute difference curve. The right panel shows that the split histogram for models trained on actual data across CVs were more similar than the null distribution generated based on permuted data (U = 0.000, $p < 0.001$, Mann-Whitney U test, Cohen's d = -3.398, $n_{actual} = 10, n_{null} = 100$). **(I)** Consistent predictions of random forest models across CV folds. To assess the output agreement, we randomly sampled an observation from the test set (in one of the folds) and passed it to two models trained from different CV folds. We repeated this procedure 200 times, both for models trained on actual data and on models trained on permuted data. We used Pearson's correlation coefficient between the predicted probabilities from a pair of models to quantify the output agreement between the two models. Each data point in the plot represents one such correlation. Models trained on actual data had higher

output agreement across CVs than those trained from permuted data (U = 40000, p < 0.001, Mann-Whitney U test, Cohen's d = 5

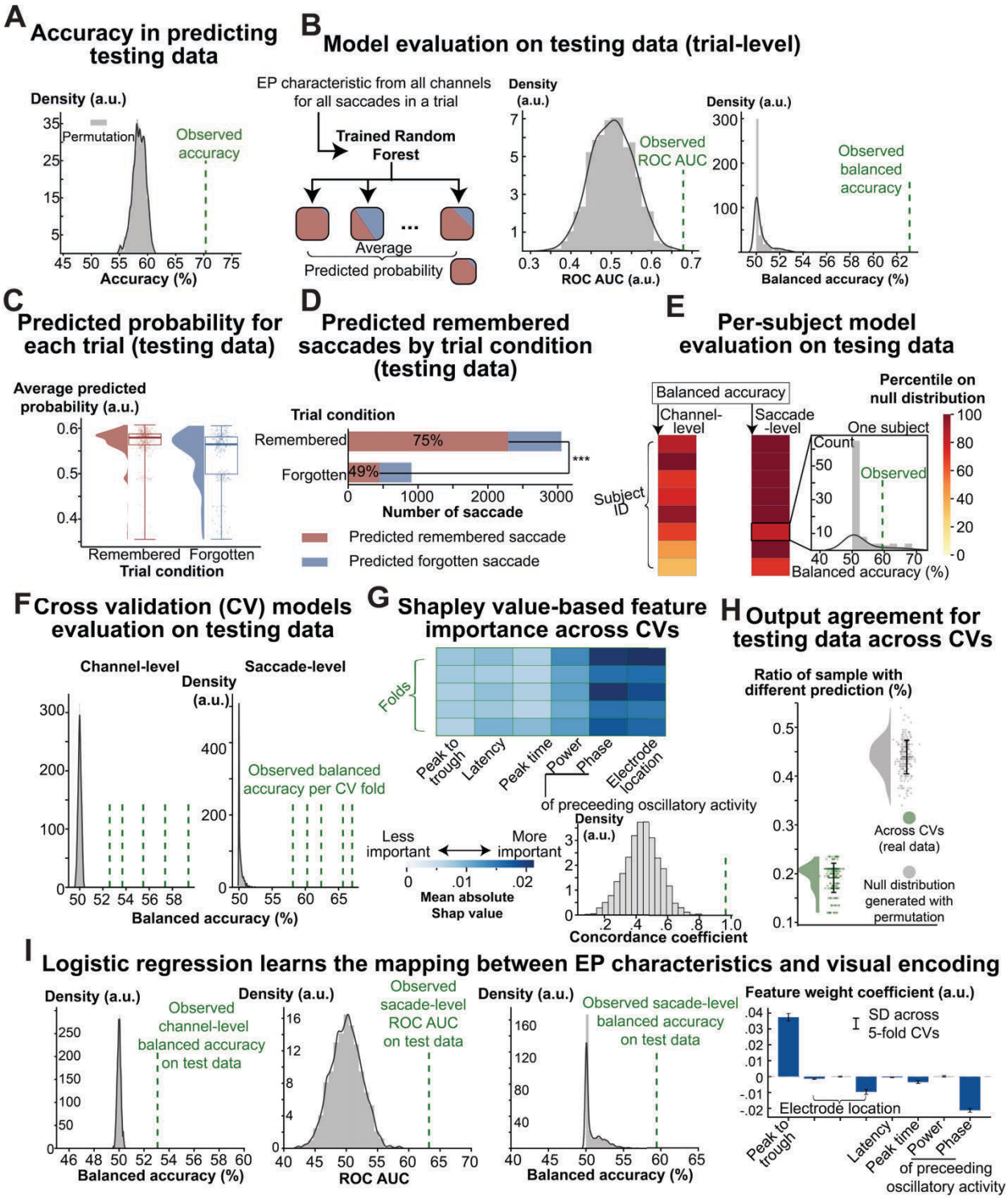


Figure 6 – Figure supplement 1. **Validation of the model's ability to learn the mapping between saccade-evoked potentials and visual encoding**. **(A)** Accuracy of

the random forest in predicting visual encoding performance on testing data. We performed 200 permutations to generate a null distribution of accuracy. The observed accuracy was 70.3%, exceeding all values in the null distribution. **(B)** Trial-level model performance on test data. We passed EP characteristics in all channels across all saccades during a trial to the model and used the average predicted probability as the trial-level predicted probability. Using these trial-level predicted probabilities, we computed the area under the receiver operating characteristic curve (ROC AUC). An illustration of calculating ROC AUC is shown in **Figure 2C**. The observed ROC AUC on test data (0.678) exceeded all values from a null distribution generated by training and evaluating models on data with permuted labels. The rightmost panel shows the observed trial-level balanced accuracy (63.1%), compared to the null distribution. To calculate the trial-level balanced accuracy, we determined the optimal threshold for average predicted probability, sweeping from 0 to 1, that maximized balanced accuracy. We applied this process to both observed data and data with permutated labels. **(C)** Average predicted probability in remembered and forgotten trials. Each data point in the plot represents the average predicted probability for a trial, and color denotes the trial condition. For test data, the average predicted probability was higher in remembered trials, compared to forgotten trials (U statistic = 84652, $p < 0.001$, rank-biserial correlation = 0.333, $n_{remembered}$ = 635, $n_{forgotten}$ = 200, Mann–Whitney U test). **(D)** Proportion of predicted remembered saccades and forgotten saccades in remembered and forgotten trials. The predicted label for a saccade was based on the average predicted probability across all recorded EPs. The trained random forest classified 75% of saccades in remembered trials as remembered saccades and 49% of saccades in forgotten trials as forgotten saccades. A chi-squared test of independence revealed a significant association between trial condition and predicted saccade labels ($\chi^2(1)$=217.89, $p < 0.001$). **(E)** Generalizability of the model across individual subjects. We evaluated the model's performance separately for each subject's test data using saccade- and channel-level balanced accuracy, averaged across cross-validation (CV) folds. For each subject, we generated a null distribution of balanced accuracy using models trained on permuted labels. The color represents the percentile rank relative to the null distribution for each subject. The observed saccade-level balanced accuracy exceeded the null distribution for 6 out of 8 subjects. We note that the other two subjects had the lowest correct recall rate. As an example, the inset shows the null distribution and observed balanced accuracy for one subject. **(F)** Channel- and saccade-level balanced accuracy of all random forest models trained in 5-fold CVs. These performance metrics exceeded all values in the null distribution. The green dashed lines show the observed balanced accuracy for each CV fold. **(G)** Concordance of Shapley value–based feature importance across CVs. We used Shapley values to interpret the contribution of each feature to the model's predictions. Shapley values, derived from cooperative game theory, quantify each feature's marginal contribution to the model's output by averaging over all possible combinations of feature

subsets. The color in the main plot represents the mean absolute Shapley value for each feature, each fold. The lower histogram shows that the Kendall's coefficient of concordance based on the mean absolute Shapley value for trained models across CVs exceeds all concordance coefficients in the null distributions. We generated the null distribution by calculating the concordance coefficient from 5 randomly sampled models trained on permuted data. **(H)** Output agreement across CVs on test data. In addition to **Figure 6I**, we assessed output agreement between models trained on different CV splits based on the ratio of samples with different predictions. For a pair of models trained on actual data in different CV folds, we passed the same test data to both models and computed the proportion of samples with different predictions. The green half violin shows the distribution of this disagreement ratio across models trained on actual data in different CV folds, and the grey half violin represents the null distribution generated from two randomly sampled models trained on permuted data. Summary bars on the right indicate the mean and standard deviation of disagreement ratios for models trained on observed or permuted labels. **(I)** Evaluation of logistic regression for mapping EP characteristics to visual encoding. As the anatomical regions and their function are not linearly distributed in the brain, we used non-linear models (random forest) to learn the mapping from EP characteristics recorded at specific brain regions to visual encoding. To assess whether there is a linear relationship, we trained logistic regression models using the same data and labels as those in **Figure 6C**. We evaluate model performance on held-out test data using channel-level balanced accuracy, saccade-level ROC AUC, and saccade-level balanced accuracy, similar to **Figure 6D–E**. We found that logistic regression performed well above chance, although not as well as the random forest model. The rightmost plot shows the weight coefficient for each feature, suggesting that a higher peak-to-trough value was associated with remembered trials. Caution is warranted when interpreting the negative coefficient for the phase of preceding oscillatory activity, as phase is a circular variable and linear models may not fully capture its periodic nature.

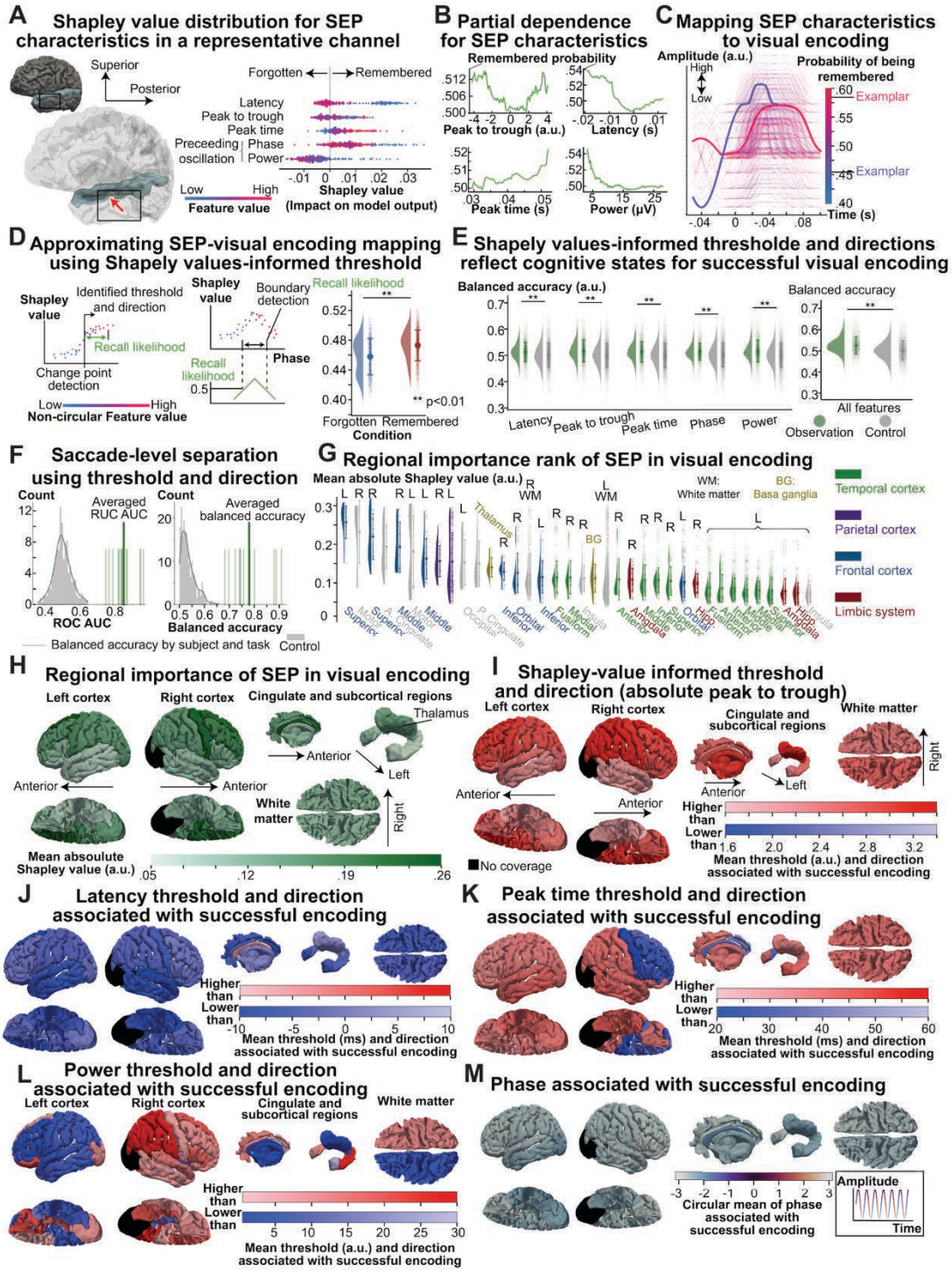


Figure 7. Explanatory analysis of the mapping between saccade-related evoked potential (EP) and visual encoding. **(A)** Interpretation of learned relationship between EP characteristics and visual encoding using Shapley values. This panel shows the Shapley values for EP characteristics recorded from a channel located at the superior temporal gyrus. Each point represents the Shapley value of a feature value, quantifying how much this feature value contributed to the model's (e.g., random forest model) prediction.

Shapley value is a method from coalitional game theory that represents the average marginal contribution of a feature value across all possible coalitions. The color represents the relative feature value, mapped separately for each feature from low to high. A positive Shapley value indicates that the feature value increased the predicted probability of successful visual encoding by the Shapley value. **(B)** Partial dependence plots for features of EPs recorded in the same channel are shown in panel **A**. Each curve represents the marginal effect of a feature on the predicted probability of successful visual encoding by averaging out the effects of all other features in the model. This panel illustrates how each feature influences a model's prediction, allowing for a better understanding of Shapley values presented in **A**. For example, in panel **A**, an earlier saccade-related EP (lower latency) is associated with higher Shapley value, which is confirmed by the partial dependence plot in panel **B.** Given that Shapley values are based on observed data and additive, we use them to interpret learned mapping in the random forest model. **(C)** Reconstructed saccade-related evoked potential and associated probability of successful encoding for the channel shown in panel **A**. We systematically varied EP characteristics, including instantaneous phase and power at half a cycle before EP latency, EP latency, peak-to-trough, and peak time. We input the feature combinations and the channel's spatial coordinates into the trained random forest model to obtain the predicted probability of successful visual encoding. Each curve represents the reconstructed EP waveform combined with preceding oscillatory activities. We reconstructed the preceding oscillatory activities using a 10Hz sinusoidal wave, with instantaneous phase and power measured at half a cycle before the EP latency. We reconstructed the EP using two sigmoid functions: one for the rising phase, beginning from the EP latency and oscillation amplitude at the EP latency, to the peak time and the peak to trough at the peak time, and one for the falling phase from the peak time onward. Reconstructed EP and preceding oscillation are color-coded by their predicted encoding probability. **(D)** Interpretation of the learned mapping in the random forest model using Shapley value-informed threshold and directions. We aimed to identify feature thresholds and directions associated with successful encoding. For non-circular features, such as absolute peak-to-trough, we ordered observations by feature values. The left plot shows the Shapley values over the ordered feature values. We detected a change point of feature values that minimized the total sum of squared differences of Shapley values from their respective mean within the segments separated by the identified change point. For the circular feature, the instantaneous phase of the preceding oscillatory activities, we searched for a window (e.g., two boundaries on the unit circle) that minimized the sum of squared Shapley value differences from the mean within and outside the window. For each feature, we defined the direction associated with successful encoding by comparing the average Shapley value between the two sides of the threshold. For the circular feature, we compared Shapley values inside and outside the identified window. To confirm that the Shapley value-informed threshold and direction explain visual encoding, we

calculated a recall likelihood, a probability-like measure of successful encoding based on the Shapley value-informed threshold and direction. For each observation, we computed recall likelihood as the distance between the feature value and the threshold, scaled by the standard deviation of the feature values. Depending on the Shapley value-informed direction, we added or subtracted the scaled distance from the prior probability (i.e., 0.5) and then clipped the result to the range [0, 1]. For circular features, the same logic was applied using the shortest angular distance between a given feature value and the boundaries of the encoding-favorable window. The right plot shows the distribution of recall likelihood averaged across channels in remembered and forgotten trials ($N_{remember} = 3052$, $N_{forgotten} = 906$, U = 1.872×10^6, p < 0.001, Mann-Whitney U test, rank-biserial correlation coefficient = 0.354). **(E)** Balanced accuracy of distinguishing between remembered and forgotten trials using recall likelihoods computed from Shapley value-informed thresholds and directions. For each saccade, we computed the recall likelihood of every EP feature at every channel and compared it against a baseline probability of 0.5. The left plot shows the distribution of balanced accuracy using actual and control Shapley value-informed thresholds and directions. We generated the control thresholds and directions by randomly sampling thresholds from the observed feature values. For all features, the balanced accuracy was higher using recall likelihood based on actual Shapley value-informed threshold and direction ($N_{observation} = 2707$, $N_{control} = 5414$, p < 0.001, t test, Cohen's d ranged from 0.296 to 0.390). The right plot shows the balanced accuracy of separating successful visual encoding from unsuccessful visual encoding using recall likelihood for all features ($N_{observation} = 2707$, $N_{control} = 5414$, U = 9.767×10^6, p < 0.001, Mann-Whitney U test, Cohen's d = 0.571). These results suggest that Shapley value-informed thresholds and directions reflect cognitive state related to successful visual encoding. We considered an observation of EP at a channel to be associated with successful visual encoding if more than half of the feature values exceeded the 0.5 baseline probability. **(F)** Saccade-level separation between remembered and forgotten trials using Shapley value-informed recall likelihood. For each saccade, we calculated the proportion of EP features associated with successful visual encoding across channels. We generated the receiver operating characteristic (ROC) curve by varying the proportion threshold. The left plot shows the distribution of area under the ROC curve (AUC), and the right plot shows the distribution of balanced accuracy at the optimal proportion threshold, comparing Shapley value-informed thresholds and directions to a control condition using thresholds randomly sampled from the feature values. Individual green lines represent the observed AUC or balanced accuracy for each subject and each task. For each plot, the thicker green line represents the averaged observed AUC (84.9%) or balanced accuracy (77.9%). The rank-biserial correlation was 1.000, and Cohen's d was 6.789 for the averaged AUC compared to the control distribution. Rank-biserial correlation was 1.000, and Cohen's d was 8.761 for balanced accuracy compared to the control distribution. **(G)** Contribution of saccade-

related EP to visual encoding across anatomical regions. To quantify the regional contribution of EPs to visual encoding, we computed the mean absolute Shapley value for each anatomical region by averaging across channels within this region and across all EP characteristics. A higher mean absolute Shapley value indicates that the EP characteristic at this region has a greater influence on visual encoding performance. **(I–M)**, Spatial distribution of EP characteristics associated with successful visual encoding across anatomical regions. The five panels visualize the Shapley value-informed direction and threshold of EP characteristics associated with successful visual encoding at each anatomical region. We used two color schemes. The red color map (light red to red) indicates regions where higher feature values were associated with successful encoding. The blue colormap (light blue to blue) indicates regions where lower feature values were associated with successful encoding. The color within the two color maps reflects the threshold value, with red and blue corresponding to more extreme thresholds. For example, a red region indicates that high feature values above a relatively high threshold are associated with successful visual encoding. For the circular EP characteristic, the instantaneous phase half a cycle before EP latency, the color denotes the angular mean of the window of phase associated with successful visual encoding. We color-coded a sinusoidal signal by phase to facilitate interpretation. The orientation of the 3D brain models matches the orientation annotated in panel **I**.

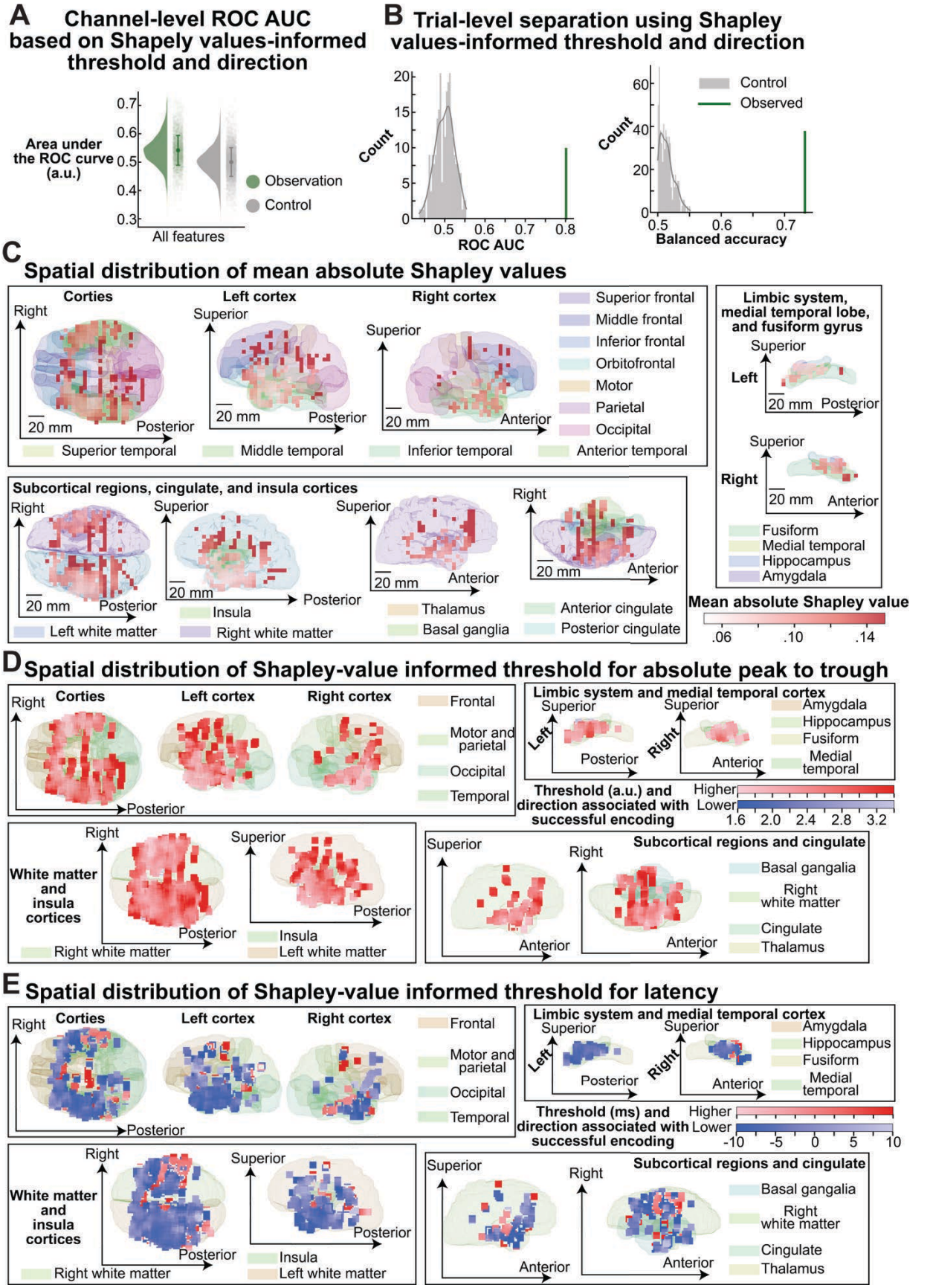

A
Channel-level ROC AUC based on Shapely values-informed threshold and direction
Area under the ROC curve (a.u.)
0.7
0.6
0.5
0.4
0.3
All features
Observation
Control
B
Trial-level separation using Shapley values-informed threshold and direction
Count
ROC AUC
Balanced accuracy
Control
Observed
C
Spatial distribution of mean absolute Shapley values
Corties
Left cortex
Right cortex
Right
Superior
Posterior
Anterior
20 mm
Superior frontal
Middle frontal
Inferior frontal
Orbitofrontal
Motor
Parietal
Occipital
Superior temporal
Middle temporal
Inferior temporal
Anterior temporal
Limbic system, medial temporal lobe, and fusiform gyrus
Left
Right
Fusiform
Medial temporal
Hippocampus
Amygdala
Subcortical regions, cingulate, and insula cortices
Insula
Thalamus
Basal ganglia
Anterior cingulate
Posterior cingulate
Left white matter
Right white matter
Mean absolute Shapley value
.06
.10
.14
D
Spatial distribution of Shapley-value informed threshold for absolute peak to trough
Frontal
Motor and parietal
Occipital
Temporal
Limbic system and medial temporal cortex
Amygdala
Hippocampus
Fusiform
Medial temporal
Threshold (a.u.) and direction associated with successful encoding
Higher
Lower
1.6
2.0
2.4
2.8
3.2
White matter and insula cortices
Subcortical regions and cingulate
Basal gangalia
Right white matter
Cingulate
Thalamus
E
Spatial distribution of Shapley-value informed threshold for latency
Threshold (ms) and direction associated with successful encoding
-10
-5
0
5
10

Figure 7 – Figure supplement 1. **Validation and illustration of Shapley value-informed thresholds and directions for successful visual encoding**. **(A)** Area under the receiver operating characteristic (ROC) curve for classifying whether a saccade-related evoked potential (EP) recorded at a channel during a saccade occurred in a remembered or forgotten trial, using recall likelihoods (detailed in **Figure 7D**) computed from Shapley value-informed thresholds and directions. We computed the ROC by systematically varying the decision threshold applied to the recall likelihood; that is, at each threshold, we computed the true positive rate and the false positive rate. We generated the control thresholds and directions by randomly sampling thresholds from the observed feature values. The area under the curve (AUC) was higher using recall likelihood based on actual Shapley value-informed threshold and direction ($N_{observation}$ = 2707, $N_{control}$ = 5414, p < 0.001, Mann-Whitney U test, U = 1.067×10^7, Cohen's d = 0.817). **(B)** Area under the ROC curve and balanced accuracy of separating remembered from forgotten trials using recall likelihoods (detailed in **Figure 7D**) aggregated across all saccades and channels for each trial. The grey distribution represents the ROC curve or balanced accuracy based on surrogate recall likelihood computed from control Shapley value-informed thresholds and directions. We generated the control thresholds and directions by randomly sampling thresholds from the observed feature values. The green vertical line represents the actual ROC curve or balanced accuracy based on actual recall likelihood. **(C)** Degree of contribution of saccade-related EP to successful visual encoding across anatomical regions. We quantified where saccade-related EPs contributed more to successful visual encoding using mean absolute Shapley values for each channel. Similar to **Figure 2 – Figure supplement 3**, we projected the 3D brain surface and electrode coordinates onto a 2D plane. We used a heatmap to represent the spatial pattern of mean absolute Sharpey values. To generate the heatmap, we discretized this 2D space into a 5 mm × 5 mm grid. For each electrode, we assigned its mean absolute Shapley value to the nearest grid element to preserve anatomical specificity. To account for the recording sensitivity of stereo-encephalogram signals, estimated to be approximately 10 mm, we computed the local average within a 10 mm radius for each grid element. Heatmap elements that have no nearby electrodes are fully transparent. Three panels show spatial distributions of mean absolute Shapley values across (i) cortices, (ii) subcortical structures, cingulate and insular cortices, and (iii) limbic regions. The heatmap color encodes the mean absolute Shapley values. The background anatomical surfaces provide spatial reference and are color-coded based on anatomical regions; their colors do not represent Shapley values. **(D)** Thresholds and directions of saccade-related EP magnitude associated with successful visual encoding across anatomical regions. To visualize how saccade-related EP magnitude relates to successful encoding, we projected the brain model and electrodes onto a 2D plane and discretized this 2D space into a 2 mm × 2 mm grid. Since directions may differ for adjacent grid elements, spatial averaging is not appropriate. To account for the recording sensitivity of

stereo-encephalogram signals, we considered electrodes within a 10 mm radius for each grid element. We determined the dominant direction, defined as whether higher or lower absolute peak-to-trough values were more frequently associated with successful encoding. For each grid element, we computed the average of Shapley value-informed thresholds for channels within a 10 mm radius whose Shapley value-informed direction matched the majority direction. We used two color schemes to represent the thresholds of saccade-related EP magnitude associated with successful visual encoding. The red color map (light red to red) indicates regions where higher feature values were associated with successful encoding. The blue colormap (light blue to blue) indicates regions where lower feature values were associated with successful encoding. Within each color map, the color intensity reflects the threshold, with darker red and blue corresponding to more extreme thresholds. **(E)** Thresholds and directions of saccade-related EP magnitude associated with successful visual encoding across anatomical regions.

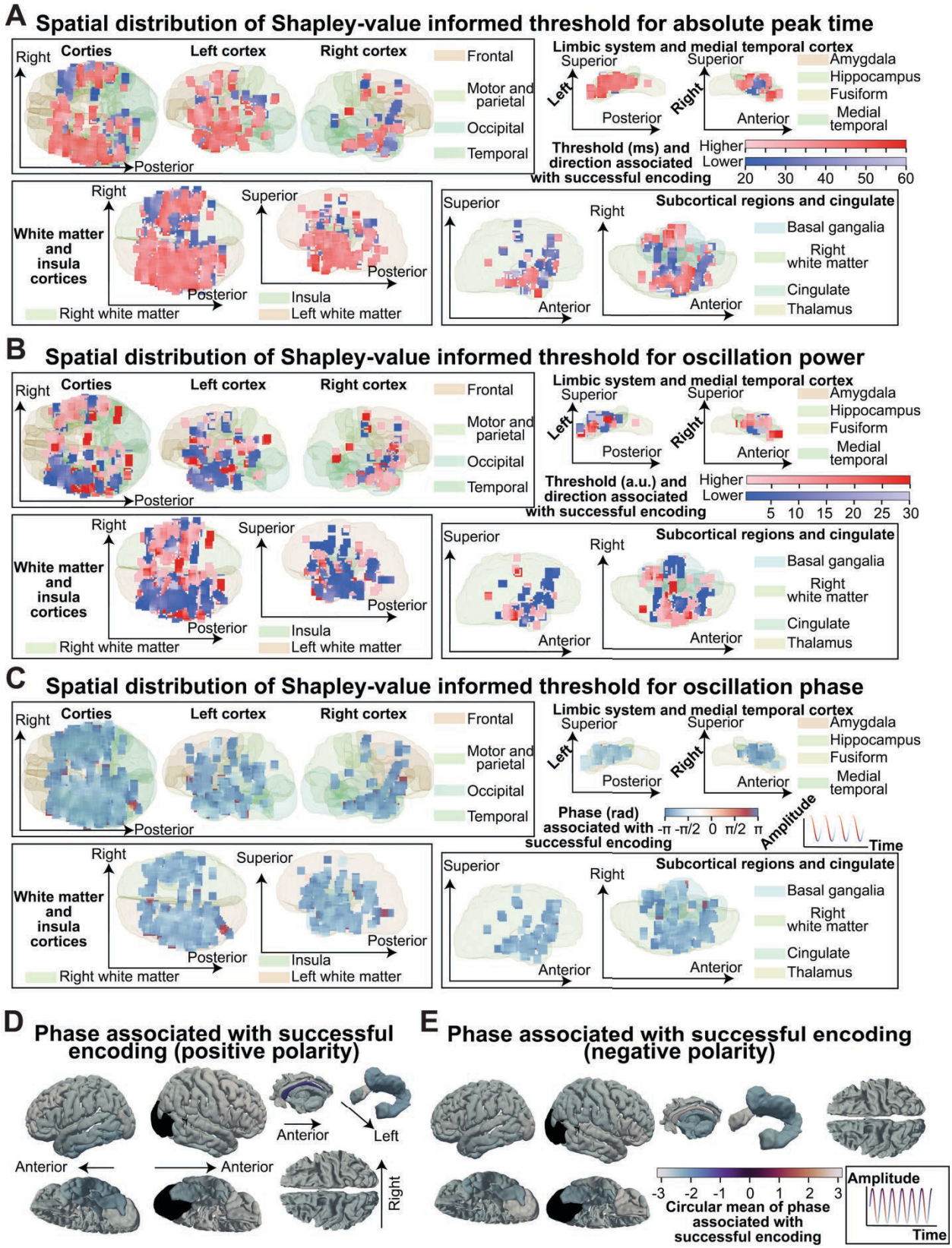

A
Spatial distribution of Shapley-value informed threshold for absolute peak time
Right
Corties
Left cortex
Right cortex
Frontal
Motor and parietal
Occipital
Temporal
Posterior
Limbic system and medial temporal cortex
Superior
Left
Posterior
Right
Anterior
Amygdala
Hippocampus
Fusiform
Medial temporal
Threshold (ms) and direction associated with successful encoding
Higher
Lower
20 30 40 50 60
White matter and insula cortices
Right
Posterior
Right white matter
Superior
Posterior
Insula
Left white matter
Subcortical regions and cingulate
Superior
Right
Anterior
Basal gangalia
Right white matter
Cingulate
Thalamus
B
Spatial distribution of Shapley-value informed threshold for oscillation power
Threshold (a.u.) and direction associated with successful encoding
5 10 15 20 25 30
C
Spatial distribution of Shapley-value informed threshold for oscillation phase
Phase (rad) associated with successful encoding
-π -π/2 0 π/2 π
Amplitude
Time
D
Phase associated with successful encoding (positive polarity)
Anterior
Left
Right
E
Phase associated with successful encoding (negative polarity)
-3 -2 -1 0 1 2 3
Circular mean of phase associated with successful encoding
Amplitude
Time

Figure 7 – Figure supplement 2. **Saccade-related EP characteristics associated with successful visual encoding across anatomical regions.** **（A–B）**. Thresholds and directions of saccade-related EP peak time and instantaneous power of the preceding oscillation associated with successful visual encoding. We computed the heatmaps using the same method as in **Figure 7 and Figure Supplement 1D**. **(C)** Instantaneous phase of the preceding oscillation associated with successful visual encoding. The heatmap color represents the angular mean of the instantaneous phase across electrodes within a 10 mm radius that were associated with successful encoding. **(D–E)**. Spatial distribution of oscillation phase half cycle before latency associated with successful visual encoding for positive (**D**) and negative polarity (**E**) EPs.

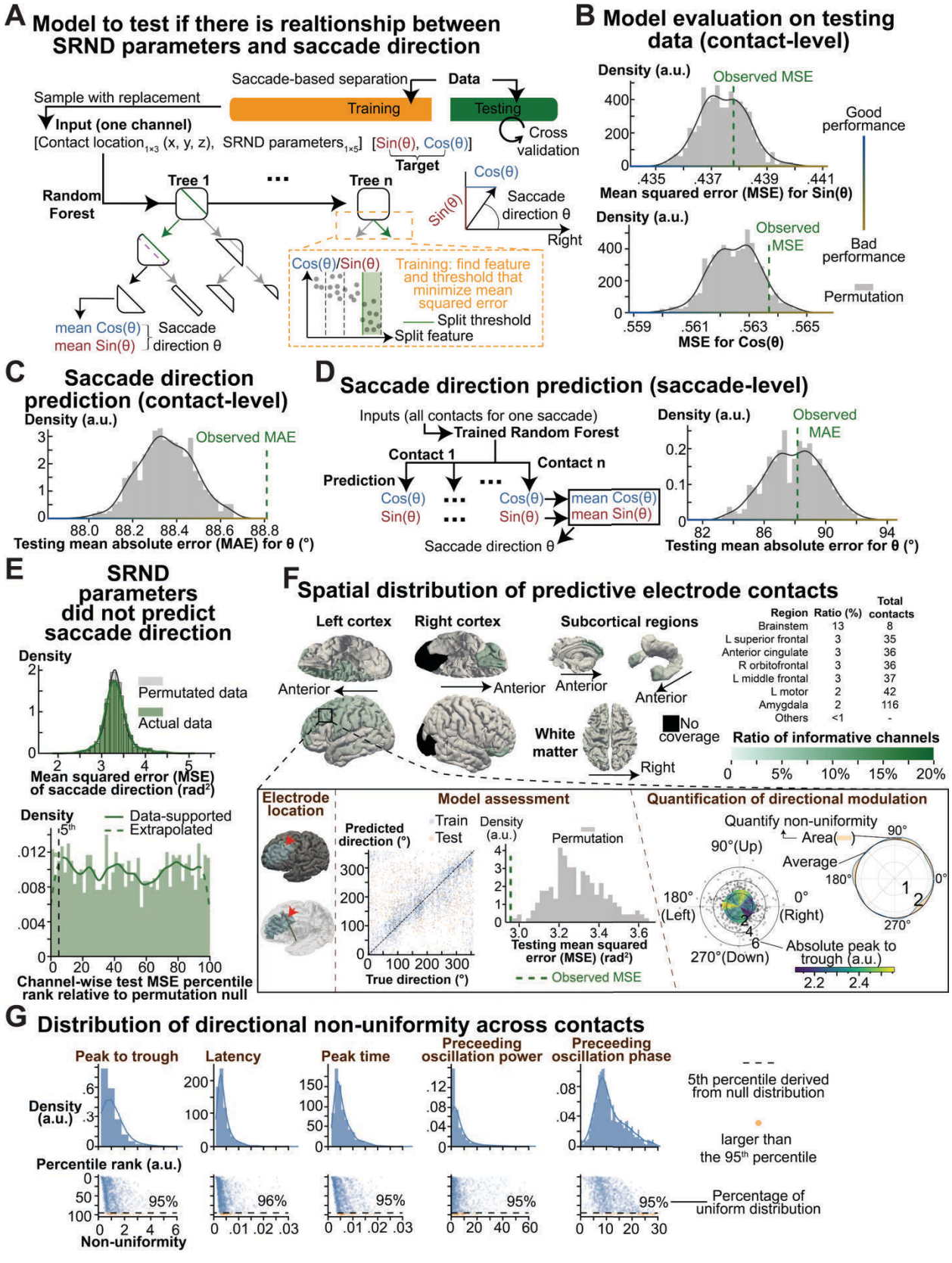

A Model to test if there is realtionship between SRND parameters and saccade direction
Saccade-based separation
Data
Sample with replacement
Training
Testing
Input (one channel)
Cross validation
[Contact location1×3 (x, y, z), SRND parameters1×5] [Sin(θ), Cos(θ)]
Target
Random Forest
Tree 1
Tree n
Cos(θ)
Sin(θ)
Saccade direction θ
Right
Cos(θ)/Sin(θ)
Training: find feature and threshold that minimize mean squared error
Split threshold
Split feature
mean Cos(θ)
mean Sin(θ)
Saccade direction θ
B Model evaluation on testing data (contact-level)
Density (a.u.)
Observed MSE
Good performance
Mean squared error (MSE) for Sin(θ)
Bad performance
Permutation
MSE for Cos(θ)
C Saccade direction prediction (contact-level)
Observed MAE
Testing mean absolute error (MAE) for θ (°)
D Saccade direction prediction (saccade-level)
Inputs (all contacts for one saccade)
Trained Random Forest
Contact 1
Contact n
Prediction
mean Cos(θ)
mean Sin(θ)
Saccade direction θ
Testing mean absolute error for θ (°)
E SRND parameters did not predict saccade direction
Density
Permutated data
Actual data
Mean squared error (MSE) of saccade direction (rad²)
Data-supported
Extrapolated
5th
Channel-wise test MSE percentile rank relative to permutation null
F Spatial distribution of predictive electrode contacts
Left cortex
Right cortex
Subcortical regions
Anterior
White matter
Right
No coverage
Region Ratio (%) Total contacts
Brainstem 13 8
L superior frontal 3 35
Anterior cingulate 3 36
R orbitofrontal 3 36
L middle frontal 3 37
L motor 2 42
Amygdala 2 116
Others <1 -
Ratio of informative channels
0 5% 10% 15% 20%
Electrode location
Model assessment
Predicted direction (°)
Train
Test
True direction (°)
Density (a.u.)
Permutation
Testing mean squared error (MSE) (rad²)
Observed MSE
Quantification of directional modulation
Quantify non-uniformity
Area
Average
90°(Up)
180° (Left)
0° (Right)
270°(Down)
Absolute peak to trough (a.u.)
G Distribution of directional non-uniformity across contacts
Peak to trough
Latency
Peak time
Preceeding oscillation power
Preceeding oscillation phase
Density (a.u.)
Percentile rank (a.u.)
95%
96%
Non-uniformity
5th percentile derived from null distribution
larger than the 95th percentile
Percentage of uniform distribution

Figure 8. **Most saccade-related evoked potentials (SREPs) did not encode saccade directions**. **(A)** Random forest regressor for learning the mapping between saccade-related neural dynamic (SRND) model parameters and saccade direction. We trained a random forest using 5-fold cross-validation. We split the training and testing data at the saccade level. The input to the random forest is a vector that includes the spatial location of electrodes and EP characteristics (e.g., latency, peak-to-trough amplitude, peak time, phase, and power of preceding oscillatory activity). To account for the circular nature of saccade direction, we defined the target as the vector $[Sin(\theta), Cos(\theta)]$, where θ denotes the saccade direction in angular space. This trigonometric transformation allows the use of Euclidean distance as a loss function and enables regression on a continuous 2D representation. During training, the random forest learned a set of features and split thresholds that minimize mean square error within each split node. To encourage spatial generalization while preserving spatial specificity, we added Gaussian-distributed spatial jitter to electrode coordinates, preventing the model from memorizing exact channel locations. After training, each leaf node in a tree outputted $[Sin(\theta), Cos(\theta)]$, averaged across the training samples that reached that node. For EP recorded in an electrode during a saccade, the predicted saccade direction is the average of the output across all trees. We obtained the predicted saccade direction using the $atan2$, two-argument arctangent function. **(B)** Model evaluation on testing data. To assess whether SRND parameters contain information about saccade direction, we evaluated the model's prediction error on held-out testing data using mean squared error (MSE) between predicted and true [sin(θ), cos(θ)] values. Lower MSE indicates a relationship between SRND parameters and saccade direction, while higher MSE indicates little such a relationship. The vertical green dashed line represents the observed MSE. The gray histogram shows the null distribution of MSE values obtained by training the model on data with permuted target directions. **(C)** Channel-level prediction of saccade direction. We quantified prediction performance at the level of individual electrodes using the mean absolute error (MAE) between predicted and true saccade directions. **(D)** Saccade-level prediction of saccade direction. For each saccade, we averaged the predicted [sin(θ), cos(θ)] vectors across electrodes to obtain the saccade-level prediction of saccade direction. The plot on the right compares the observed MAE against a null distribution generated by permuting saccade directions during training. **(E)** SRND parameters in most electrode contacts did not contain information about saccade direction. The top plot shows the distribution of MSE between predicted and true saccade directions across electrode contacts, compared to a null distribution. For each electrode contacts, we computed the observed MSE and generated a null distribution by permuting saccade directions during training 200 times. The bottom plot shows the distribution of MSE percentile rank relative to each channel's null distribution. The vertical black dashed line marks the 5th percentile, and the green dashed line represents kernel density estimation outside the data support. **(F)** Ratio of saccade direction-informative channels across

anatomical regions. We identified channels where EP characteristics were informative of saccade direction by comparing the testing MSE with the 5th percentile of the null distribution, adjusted for the total number of channels. The color of each anatomical region represents the proportion of saccade direction-informative channels among all channels within that region. The inset illustrates an example informative electrode located in the left middle frontal lobe. In the two 3D brain models on the left, the left middle frontal lobe is highlighted in blue, and red arrows indicate the electrode location. The first middle plot in the inset illustrates the relationship between the predicted and true saccade directions for training and testing data. The second middle plot shows the observed MSE for the testing data and the null distribution. The rightmost plots show the explanatory analysis of directional modulation for each EP characteristic. For example, we estimated the circular kernel density of the absolute peak-to-trough value as a function of saccade direction. We quantified directional modulation as non-uniformity, operationally measured as the integral of the absolute difference between the kernel density curve and the mean absolute peak-to-trough value across the angular domain. **(G)** Distribution of directional non-uniformity across channels. The top row shows the distribution of non-uniformity for SRND parameters. The bottom row shows the corresponding percentile rank of non-uniformity relative to each channel's null distribution. Each data point represents a channel. The horizontal dashed line marks the 5th percentile of the null distribution. The annotated percentage indicates the proportion of channels whose non-uniformity is less than the 95th percentile of their relative null distribution.

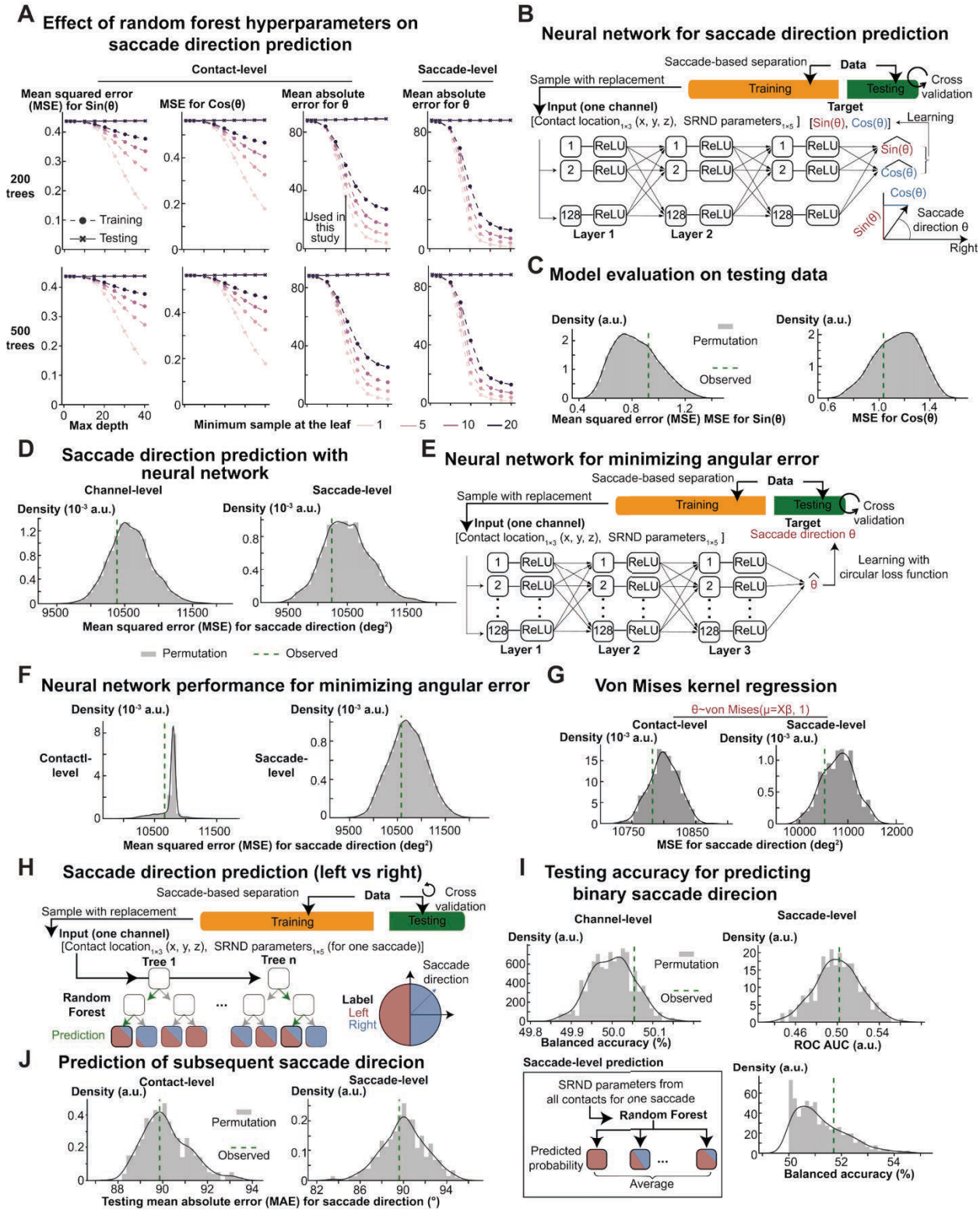
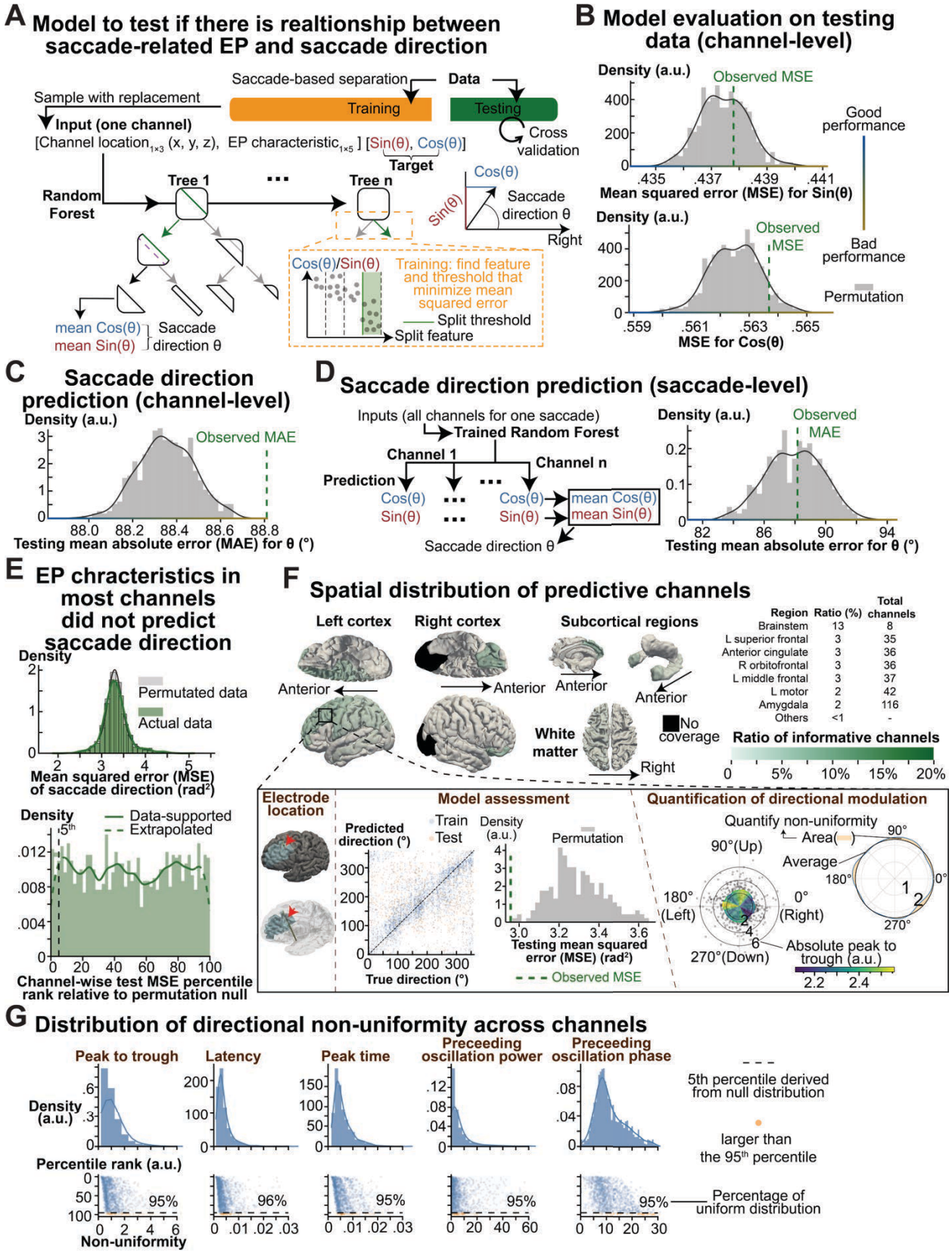


Figure 8 – Figure supplement 1. **Most SREPs did not encode saccade directions, nor subsequent saccade directions. (A)** Effect of hyperparameters on model performance during training and testing. We evaluated how different hyperparameter

settings affected the performance of random forests for predicting saccade direction. The hyperparameter includes the number of trees, the maximum tree depth, and the minimum sample size at the leaf. In general, deeper trees with a small sample size can learn more complex patterns but risk overfitting. In each plot, color represents the minimum sample at the leaf, and the x-axis represents the maximum depth. From left to right, the y-axis represents channel-level mean squared error (MSE) for $Sin(\theta)$, $Cos(\theta)$, saccade direction θ, and saccade-level MSE for saccade direction θ The top row presents the model performance using 200 decision trees, and the bottom row shows the model performance with 500 decision trees. In this study, we used a default hyperparameter setting of 200 decision trees, a maximum depth of 20, and a minimum leaf size of 1. **(B)** Neural network architecture for predicting saccade direction. To rule out the possibility that the random forest model might not capture the relationship between saccade direction and SRND parameters, we implemented a neural network with three fully connected hidden layers. The input included electrode contact locations and EP characteristics. The output layer had two units representing sine and cosine of the saccade direction $Sin(\theta)$, $Cos(\theta)$. Each unit in the hidden layers was followed by a Rectified Linear Unit (ReLU) activation function to model non-linearity. **(C)** Evaluation of the neural network for predicting saccade direction. The two plots show the observed channel-level mean squared error (MSE) for $Sin(\theta)$ and $Cos(\theta)$, compared to null distributions of MSE obtained from neural networks trained with permuted data. **(D)** Electrode conact- and saccade-level MSE for saccade direction θ. To calculate saccade-level MSE, we averaged the outputs [$Sin(\theta)$, $Cos(\theta)$] across channels, and computed the predicted saccade direction using $atan2$, the two-argument arctangent function. The vertical green dashed marks the observed MSE for saccade direction, and the grey histogram represents the null distribution of MSE obtained from neural networks trained with permuted data. **(E)** Neural network architecture for minimizing angular loss. As a complementary model to predicting the trigonometric function of the saccade angle (shown in **B**), we implemented a neural network to minimize angular loss. There is one unit at the output layer representing the saccade direction. **(F)** Evaluation of a neural network for minimizing angular loss. The vertical green dashed lines show the observed channel- and saccade-level MSE between predicted and actual saccade direction. The grey histogram represents the null distribution for MSE, which was generated from neural networks trained with permuted data. **(G)**. Saccade direction prediction using circular kernel (i.e., von Mises kernel) regression. To assess if SRND parameters linearly contribute to saccade direction, we fitted a GLM with a von Mise link function. The plots show the observed channel- and saccade-level MSE between predicted and actual saccade direction and their null distribution. **(H)** Random forest model for classifying binary saccade direction using SRND parameters. We binarized saccade direction as left vs. right saccade, comparable to the prediction of successful visual encoding (**Figure 6C**). We trained the random forest classifier with inputs of SRND parameters. **(I)** Binary classification of saccade direction

using random forest classifier. The top-left plot shows the observed electrode contact-level balanced accuracy for testing data and the null distribution of testing balanced accuracy. We generated the null distribution of testing balanced accuracy by training the random forest classifier with permuted binary labels of saccade direction. Similar to visual encoding prediction, we used the averaged predicted probability for rightward saccades across channels during a saccade as the saccade-level predicted probability. The right plots show the distribution of saccade-level area under the receiver operating characteristic (ROC) curve and the optimal balanced accuracy. We obtained the threshold for optimal balanced accuracy by varying the threshold from 0 to 1. We generated the null distribution of saccade-level area under the ROC curve and balanced accuracy using the same procedure but with permuted data. **(J)** Prediction of subsequent saccade direction using random forest regressor. We identified subsequent saccades as the next saccade within 500ms. The random forest regressor structure is shown in **Figure 8A**. The two plots show the observed channel- and saccade-level testing mean absolute error for saccade direction. The grey histogram represents the null distribution of observed channel- and saccade-level testing mean absolute error, generated from models trained with permuted data.



Region	Ratio (%)	Total channels
Brainstem	13	8
L superior frontal	3	35
Anterior cingulate	3	36
R orbitofrontal	3	36
L middle frontal	3	37
L motor	2	42
Amygdala	2	116
Others	<1	-

Figure 8. **Most saccade-related evoked potentials (EPs) did not encode saccade directions**. **(A)** Random forest regressor for learning the mapping between saccade-related EP and saccade direction. We trained a random forest using 5-fold cross-validation. We split the training and testing data at the saccade level. The input to the random forest is a vector that includes the spatial location of electrodes and EP characteristics (e.g., latency, peak-to-trough amplitude, peak time, phase, and power of preceding oscillatory activity). To account for the circular nature of saccade direction, we defined the target as the vector [$Sin(\theta)$, $Cos(\theta)$], where θ denotes the saccade direction in angular space. This trigonometric transformation allows the use of Euclidean distance as a loss function and enables regression on a continuous 2D representation. During training, the random forest learned a set of features and split thresholds that minimize mean square error within each split node. To encourage spatial generalization while preserving spatial specificity, we added Gaussian-distributed spatial jitter to electrode coordinates, preventing the model from memorizing exact channel locations. After training, each leaf node in a tree outputted [$Sin(\theta)$, $Cos(\theta)$], averaged across the training samples that reached that node. For EP recorded in an electrode during a saccade, the predicted saccade direction is the average of the output across all trees. We obtained the predicted saccade direction using the $atan2$, two-argument arctangent function. **(B)** Model evaluation on testing data. To assess whether saccade-related EPs contain information about saccade direction, we evaluated the model's prediction error on held-out testing data using mean squared error (MSE) between predicted and true [sin(θ), cos(θ)] values. Lower MSE indicates a relationship between saccade-related EPs and saccade direction, while higher MSE indicates little such a relationship. The vertical green dashed line represents the observed MSE. The gray histogram shows the null distribution of MSE values obtained by training the model on data with permuted target directions. **(C)** Channel-level prediction of saccade direction. We quantified prediction performance at the level of individual electrodes using the mean absolute error (MAE) between predicted and true saccade directions. **(D)** Saccade-level prediction of saccade direction. For each saccade, we averaged the predicted [sin(θ), cos(θ)] vectors across electrodes to obtain the saccade-level prediction of saccade direction. The plot on the right compares the observed MAE against a null distribution generated by permuting saccade directions during training. **(E)** Saccade-related EP characteristics in most channels did not contain information about saccade direction. The top plot shows the distribution of MSE between predicted and true saccade directions across channels, compared to a null distribution. For each channel, we computed the observed MSE and generated a null distribution by permuting saccade directions during training 200 times. The bottom plot shows the distribution of MSE percentile rank relative to each channel's null distribution. The vertical black dashed line marks the 5th percentile, and the green dashed line represents kernel density estimation outside the data support. **(F)** Ratio of saccade direction-informative channels across anatomical regions. We identified channels where EP characteristics

were informative of saccade direction by comparing the testing MSE with the 5th percentile of the null distribution, adjusted for the total number of channels. The color of each anatomical region represents the proportion of saccade direction-informative channels among all channels within that region. The inset illustrates an example informative electrode located in the left middle frontal lobe. In the two 3D brain models on the left, the left middle frontal lobe is highlighted in blue, and red arrows indicate the electrode location. The first middle plot in the inset illustrates the relationship between the predicted and true saccade directions for training and testing data. The second middle plot shows the observed MSE for the testing data and the null distribution. The rightmost plots show the explanatory analysis of directional modulation for each EP characteristic. For example, we estimated the circular kernel density of the absolute peak-to-trough value as a function of saccade direction. We quantified directional modulation as non-uniformity, operationally measured as the integral of the absolute difference between the kernel density curve and the mean absolute peak-to-trough value across the angular domain. **(G)** Distribution of directional non-uniformity across channels. The top row shows the distribution of non-uniformity for saccade-related EP characteristics. The bottom row shows the corresponding percentile rank of non-uniformity relative to each channel's null distribution. Each data point represents a channel. The horizontal dashed line marks the 5th percentile of the null distribution. The annotated percentage indicates the proportion of channels whose non-uniformity is less than the 95th percentile of their relative null distribution.

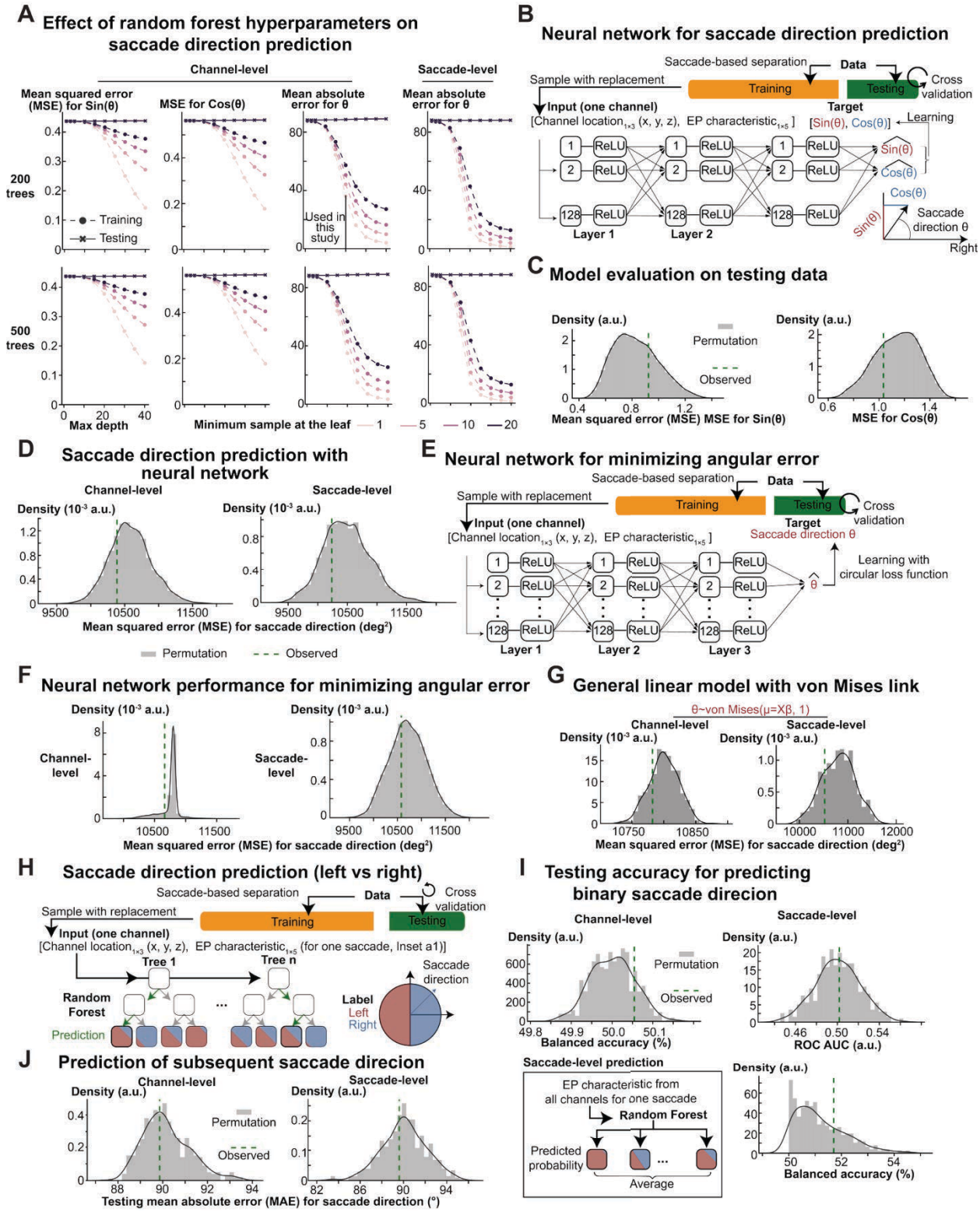


Figure 8 – Figure supplement 1. **Most saccade-related evoked potentials (EPs) did not encode saccade directions, nor subsequent saccade directions. (A)** Effect of hyperparameters on model performance during training and testing. We evaluated how

different hyperparameter settings affected the performance of random forests for predicting saccade direction. The hyperparameter includes the number of trees, the maximum tree depth, and the minimum sample size at the leaf. In general, deeper trees with a small sample size can learn more complex patterns but risk overfitting. In each plot, color represents the minimum sample at the leaf, and the x-axis represents the maximum depth. From left to right, the y-axis represents channel-level mean squared error (MSE) for $Sin(\theta)$, $Cos(\theta)$, saccade direction θ, and saccade-level MSE for saccade direction θ The top row presents the model performance using 200 decision trees, and the bottom row shows the model performance with 500 decision trees. In this study, we used a default hyperparameter setting of 200 decision trees, a maximum depth of 20, and a minimum leaf size of 1. **(B)** Neural network architecture for predicting saccade direction. To rule out the possibility that the random forest model might not capture the relationship between saccade direction and saccade-related EP characteristics, we implemented a neural network with three fully connected hidden layers. The input included electrode contact locations and EP characteristics. The output layer had two units representing sine and cosine of the saccade direction $Sin(\theta)$, $Cos(\theta)$. Each unit in the hidden layers was followed by a Rectified Linear Unit (ReLU) activation function to model non-linearity. **(C)** Evaluation of the neural network for predicting saccade direction. The two plots show the observed channel-level mean squared error (MSE) for $Sin(\theta)$ and $Cos(\theta)$, compared to null distributions of MSE obtained from neural networks trained with permuted data. **(D)** Channel- and saccade-level MSE for saccade direction θ. To calculate saccade-level MSE, we averaged the outputs [$Sin(\theta)$, $Cos(\theta)$] across channels, and computed the predicted saccade direction using $atan2$, the two-argument arctangent function. The vertical green dashed marks the observed MSE for saccade direction, and the grey histogram represents the null distribution of MSE obtained from neural networks trained with permuted data. **(E)** Neural network architecture for minimizing angular loss. As a complementary model to predicting the trigonometric function of the saccade angle (shown in **B**), we implemented a neural network to minimize angular loss. There is one unit at the output layer representing the saccade direction. **(F)** Evaluation of a neural network for minimizing angular loss. The vertical green dashed lines show the observed channel- and saccade-level MSE between predicted and actual saccade direction. The grey histogram represents the null distribution for MSE, which was generated from neural networks trained with permuted data. **(G)**. Predicting saccade direction using a generalized linear model (GLM) with von Mises link function. To assess if saccade-related EP characteristics linearly contribute to saccade direction, we fitted a GLM with a von Mise link function. The plots show the observed channel- and saccade-level MSE between predicted and actual saccade direction and their null distribution. **(H)** Random forest model for classifying binary saccade direction using saccade-related EP characteristics. We binarized saccade direction as left vs. right saccade, comparable to the prediction of successful visual encoding (**Figure 6C**). We trained the random forest

classifier with inputs of saccade-related EP characteristics. **(I)** Random forest classifier cannot separate left saccades from right saccades. The top-left plot shows the observed channel-level balanced accuracy for testing data and the null distribution of testing balanced accuracy. We generated the null distribution of testing balanced accuracy by training the random forest classifier with permuted binary labels of saccade direction. Similar to visual encoding prediction, we used the averaged predicted probability for rightward saccades across channels during a saccade as the saccade-level predicted probability. The right plots show the distribution of saccade-level area under the receiver operating characteristic (ROC) curve and the optimal balanced accuracy. We obtained the threshold for optimal balanced accuracy by varying the threshold from 0 to 1. We generated the null distribution of saccade-level area under the ROC curve and balanced accuracy using the same procedure but with permuted data. **(J)** Most saccade-related EPs did not encode subsequent saccade directions. We used a random forest regressor to predict subsequent saccade direction. We identified subsequent saccades as the next saccade within 500ms. The random forest regressor structure is shown in **Figure 8A**. The two plots show the observed channel- and saccade-level testing mean absolute error for saccade direction. The grey histogram represents the null distribution of observed channel- and saccade-level testing mean absolute error, generated from models trained with permuted data.