Statistical learning drives anticipatory micro-saccades toward suppressed distractor locations

Sirui Chen[1,2], Xin Zhang[1,2,3], Xinyu Li[3], Ole Jensen[4], Jan Theeuwes[5], and Benchi Wang[1,2*]

[1] Center for Studies of Psychological Application, South China Normal University, Guangzhou, China
[2] Institute for Brain Research and Rehabilitation, South China Normal University, Guangzhou, China
[3] Department of Psychology, Zhejiang Normal University, Jinhua, China
[4] Department of Psychiatry and Experimental Psychology, University of Oxford, Oxford, United Kingdom
[5] Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

Author notes

BW designed the experiment, SC, XZ, and XL collected and analyzed the data. SC, OJ, JT and BW wrote the paper. All authors approved the final version of the manuscript for submission. All authors declared no competing interests. This research was supported by the MOE Project of Key Research Institute of Humanities and Social Sciences in Universities (22JJD190006) and the Natural Science Foundation of Guangdong grant (2023A1515012789) to BW; the Welcome Trust Discovery Award (227420) and the NIHR Oxford Health Biomedical Research Centre (203316) to OJ. Correspondence should be addressed to Benchi Wang, Institute for Brain Research and Rehabilitation, South China Normal University, Zhongshan Road West 55, Guangzhou, China, 510000. Email: wangbenchi.swift@gmail.com. The processing codes and data will be made available upon reasonable request.

Abstract

Statistical learning enables individuals to suppress distracting but salient information, yet the mechanisms underlying this suppression remain unclear. While some suggest that suppression is proactive, occurring without first attending to distractor locations, others argue that it is reactive, requiring covert attention toward the distractor location before suppression can occur. To address this fierce debate, the current study recorded micro-saccades—an index of covert attention—during the pre-stimulus interval, alongside electroencephalogram (EEG) data collected during a visual search task. Participants were instructed to ignore a salient distractor that appeared more frequently at one specific location than at others. Learning high-probability distractor locations reduced attentional capture by salient distractors, coinciding with increased micro-saccade rates and decodable representations of these locations in alpha power (8-14 Hz) before stimulus onset. Strikingly, these anticipatory micro-saccades were more frequently directed toward high-probability distractor locations than away from them, supporting a reactive suppression mechanism. Overall, these findings highlight a crucial role for the oculomotor system in encoding and responding to learned spatial regularities.

Keywords: Statistical learning, distractor suppression, micro-saccades, alpha suppression

Introduction

The ability to extract statistical regularities (i.e., statistical learning) from the environment is fundamental to human survival. Recent studies have highlighted the role of statistical learning on attentional selection[1–3], as well as the suppression of distracting salient stimuli[4,5]. For instance, distractors appearing frequently at specific locations cause less interference than those appearing at other locations, indicating that implicitly learned regularities facilitate distractor suppression[6–11]. This learned suppression likely enhances perceptual efficiency and supports action planning when navigating in complex environments.

Despite its importance, the neural mechanism underlying learned suppression remains poorly understood, with ongoing debate regarding the nature of attentional suppression[12]. The suppression can, in principle, be implemented in two ways: proactively (before the display appears) or reactively (following attentional engagement). Studies on statistical learning propose that the spatial priority map is proactively altered, reducing competition for attention at distractor locations[3,13–18]. This proactive adjustment diminishes the distractor saliency and minimizes their attentional capture over sustained periods. Conversely, other studies provide evidence for reactive distractor suppression, where suppression occurs only after attention has been engaged[19–26], resembling the classic phenomenon of inhibition of return[27,28].

Notably, the human oculomotor system might offer a compelling avenue for investigating this issue, given its close link to covert spatial attention, as evidenced by micro-saccades. It is well-established that micro-saccades tend to occur in the direction of covert attention[29–32]. Furthermore, micro-saccades occur not only when directing attention to external locations[29,33], but also when directing attention to internal memory representations by retro-cues[32,34]. This highlights the role of micro-saccades in reflecting the allocation of covert attention to specific locations or memory-based representations, even in the absence of visual stimuli. These findings provide a unique method, by measuring micro-saccades, for examining whether

covert attention precedes attentional suppression (supporting reactive mechanism) or suppression can occur independently of prior attention allocation (supporting proactive mechanism).

Thus, in the current study, we measured micro-saccades during the pre-stimulus interval, assessing the involvement of covert attention prior to display onset, and examined its relationship with neural activity originating in the dorsal attention network. We simultaneously recorded electroencephalogram (EEG) and eye movements while participants performed the additional singleton task[6,35]. In this task, participants were instructed to search for a unique shape (the target) while ignoring a salient distractor colored differently (see Fig. 1a). The salient distractor was presented more frequently at one specific location (high-probability location) compared to others (low-probability location; see Fig. 1b and Methods for details), introducing a statistical learning component, resulting in learned suppression of the salient distractor.

Results

Mean response times (RTs) were slower when distractors were present compared to absent (755 ms), regardless of whether they were present in high-probability locations (825 ms), $t(26) = 8.84$, $p < .001$, or low-probability locations (873 ms), $t(26) = 10.02$, $p < .001$ (Fig. 1c), reflecting capture by salient distractors[35,36]. Notably, faster responses were observed when the salient distractor was presented at the high- compared to low-probability locations, $t(26) = 7.87$, $p < .001$ (Fig. 1c) suggesting suppression of the high-probability distractor location (c.f., refs.[6]).

Anticipatory micro-saccades during pre-stimulus gazes

As hypothesized, learned suppression may engage the oculomotor system before stimulus onset. To test this, we examined the micro-saccade rate during the fixation period preceding the search array onset. As shown in Fig. 1d, micro-saccade rates exhibited a rapid increase after fixation presentation followed by a gradual decline after peaking. When comparing with the average rate across time, we found

significant increases in micro-saccade rates during the interval of -2370 to -1136 ms relative to the stimulus onset (cluster-based permutation test, $p < .05$). The elevated micro-saccade rate during the pre-stimulus interval suggests that covert attention was engaged in preparation for the search task.

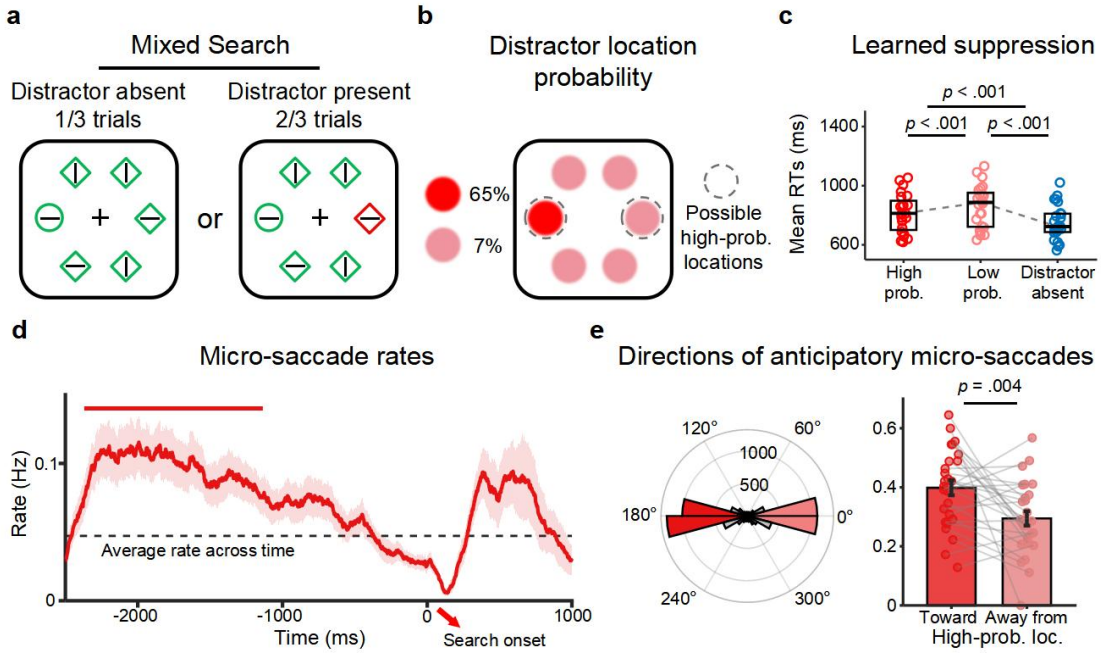

Figure 1. a) Examples of the search display. Participants searched for a unique shape (e.g., the circle on the left) and reported the orientation of the line inside it (e.g., horizontal), while ignoring a salient distractor (e.g., the red diamond). Distractor present and absent trials were randomly intermixed within each block. b) Probabilities of distractor location. The salient distractor was presented either on the left or right side (counterbalanced between participants) with a high probability of 65%, while was presented in the other five locations with an equal probability of 7% each. c) Behavioral results, with individual RTs represented by dots. High prob. and Low prob. refer to high- and low probability distractor locations. d) Micro-saccade rates, with data variability represented by ±1 SEM. Time windows, showing that micro-saccade rates were significantly larger than their average rate across time are highlighted by solid grey lines (cluster-based permutation test, $p < .05$). e) The left panel shows polar-histograms of anticipatory micro-saccades directions, and the right panel shows ratios of micro-saccades directed toward or away from high-probability distractor locations

Critically, a key finding was that a greater proportion of micro-saccades during the fixation period were directed toward the high-probability distractor location (39.6%) compared to those directed away (28.5%), $t(26) = 3.16$, $p = .004$ (Fig. 1e). This indicates that covert attention was proactively biased toward high-probability location, even though the behavioral data showed that this location was subsequently

suppressed.

Tracking learned high-probability locations from pre-stimulus alpha oscillation

Previous studies have identified contralateral pre-stimulus alpha oscillations over the parietal and occipital cortex associated with statistically learned suppression at high-probability locations[16]. Yet, in the present study, we did not observe such lateralized modulation (see Fig. 2a). Instead, pre-stimulus alpha power over both contralateral and ipsilateral electrodes relative to the high-probability location showed a sustained increase, spanning from -2346 ms to -456 ms and -2340 ms to -414 ms before the stimulus onset, respectively. This was followed by a reduction in alpha power from 232 ms to 1000 ms and from 230 ms to 1000 ms after the onset of the search, respectively (cluster-based permutation test, $p < .05$; Fig. 2a). To determine whether the observed alpha oscillation reflected attention being directed toward or away from the high-probability location, we employed a multi-variate inverted encoding model (IEM, refs.[37]; see Methods for details), which is capable of decoding neuronal representations associated with direction of spatial attention from the brain activity distributed across parietal and occipital electrodes. We successfully tracked spatial representation of the high-probability location from alpha power during the -320 – 20 ms interval relative to the stimulus onset (cluster-based permutation test, $p < .05$; Fig. 2b). No such representation was observed for low-probability locations. These findings suggest that the spatial distribution of pre-stimulus alpha oscillation may reflect enhanced feedforward control mechanisms over high-probability locations, establishing a preparatory state that anticipates and encodes the spatial representation of the high-probability location.

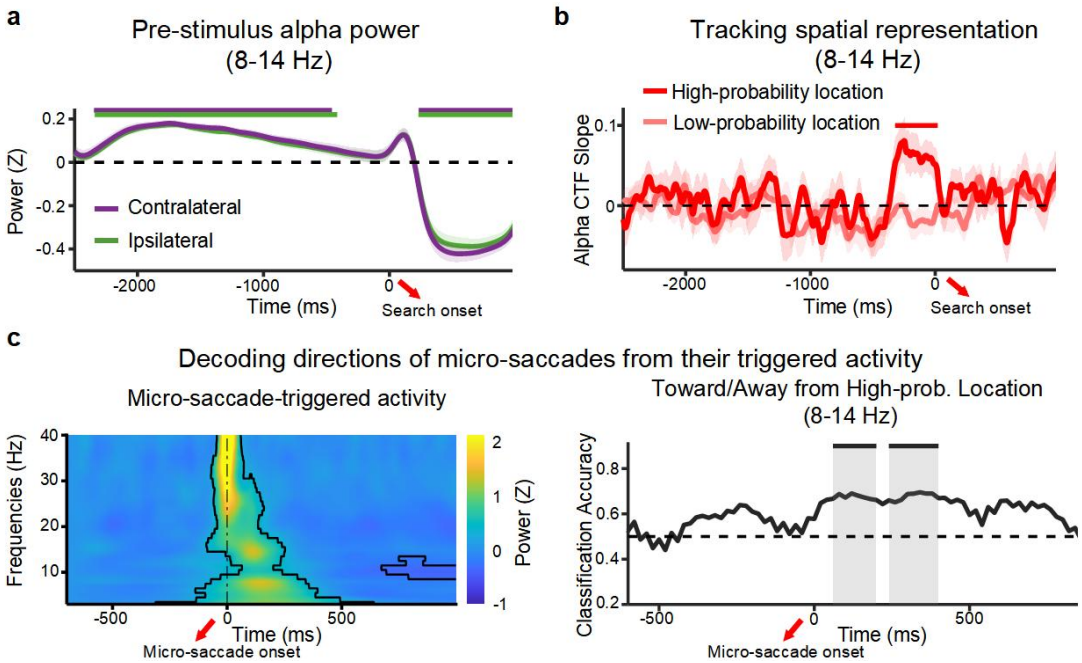

Figure 2. a) Pre-stimulus alpha oscillation over the contralateral and ipsilateral electrodes relative to the high-probability distractor location (cluster-based permutation test, $p < .05$). b) The spatial representations of high- and low-probability distractor locations were tracked within alpha oscillations (8-14 Hz) using an inverted encoding model (see Methods for details). Significant CTF slopes are highlighted by solid lines (cluster-based permutation test, $p < .05$). c) The time-frequency power spectrum of micro-saccade-triggered activities (left panel), and the accuracy for classifying anticipatory micro-saccade directed toward versus away from the high-probability location based on micro-saccade-triggered alpha activity (right panel).

Micro-saccade-triggered activity

To further validate the role of micro-saccades in driving neural activity associated with learned suppression, we investigated whether micro-saccade-triggered brain activity across multiple frequency bands encoded information about learned statistical regularities. Previous studies have shown that micro-saccades elicit posterior neural activity detectable in EEG signals[38–41], linking them closely to attentional processing. This association suggests that micro-saccade-triggered brain activity may also be sensitive to learned suppression, as such suppression involves attention-based learning. To test this hypothesis, we analyzed the power of micro-saccade-triggered neural oscillations over occipital-parietal cortices across multiple frequency bands (3-40 Hz; see Methods for details). The results showed a general increase in power across in a

wide range of frequencies (3-40 Hz) following micro-saccade onset (Fig. 2c, left panel). When isolating alpha activity (8-14 Hz) from raw EEG signals, we found evidence that micro-saccade-triggered alpha activity encoded learned statistical regularities, tracking the direction of anticipatory micro-saccades. Specifically, using a multivariate decoding approach (see Methods for details), we successfully discriminated micro-saccades directed toward and away from the high-probability location based on micro-saccade-triggered alpha activity during intervals of 60 - 200 ms and 240 - 400 ms after the anticipatory micro-saccade onset (cluster-based permutation test, $p < .05$; Fig. 2c, right panel). Notably, no significant alpha activity was observed in the micro-saccade-triggered signals recorded from electro-oculogram (EOG) electrodes, ruling out the possibility that the observed findings were driven by artifacts associated with micro-saccades[38]. These results suggest that, although micro-saccades may not be obligatory for task completion, they indeed play a role in triggering neural activity that encodes and processes learned information.

Discussion

The current study shows that learned suppression proactively recruits the oculomotor system. Consistent with previous studies, we observed the classic attentional capture effect[35,36] and reduced capture for distractors presented at the high-probability location, indicating learned suppression[3,4,6–8,10]. A novel finding of our study is the increased micro-saccade rate observed before the search display onset (Fig. 1d), indicating the involvement of covert attention in anticipating the arrival of the search display. Critically, and perhaps unexpectedly, these anticipatory micro-saccades were more frequently directed toward high-probability distractor locations than away from them (Fig. 1e). Consistent with this finding, we successfully tracked high-probability locations in pre-stimulus alpha power (8-14 Hz). Further analysis revealed that micro-saccades could trigger neural activity sensitive to learned regularities, with micro-saccade-triggered alpha activity encoding and tracking the direction of anticipatory micro-saccades (Fig. 2c). Collectively, these findings suggest

that learned suppression engages the oculomotor system, potentially functioning as a monitoring mechanism for locations likely to contain distractors.

Micro-saccades are deeply integrated into various cognitive processes, including attentional control[29,33], temporal expectation[42–44], working memory[32,34], and perceptual learning[45]. Our study extends these findings by showing that anticipatory micro-saccades are actively involved in learned suppression associated with statistical regularities (i.e., high-probability distractor locations). While previous studies have claimed that learned suppression can occur proactively without covert attention first being directed toward the suppressed location[3,13,15–18,46,47], the current findings suggest otherwise. We speculate that attending to the distractor location proactively might provide an advantage, as it allows the observer to eliminate it as a potential target location. This process triggers the fast disengagement of covert attention from the distractor location when the search array is presented, ultimately resulting in learned suppression. These findings align with the “search and destroy” hypothesis[24], which claims that suppression is only possible after first attending to the location that needs to be ignored.

Furthermore, from a theoretical perspective, the involvement of the oculomotor system in learned suppression could be explained by its role in processing spatial information. As the brain learns to identify and track locations likely to contain distractors, it may rely on the oculomotor system's precise spatial representation capabilities. This idea aligns with previous research showing that remembering spatial locations activates the oculomotor system, leading to eye movements that curve away from locations marked for later recall[48,49].

Previous studies have identified pre-stimulus alpha oscillations over the parietal and occipital cortices associated with statistically learned suppression[16]. Specifically, increased parieto-occipital alpha power contralateral to high-probability distractor locations has been observed up to one second before stimulus onset, consistent with inhibitory processes linked to distractor suppression[50]. However, in the present study, we did not observe such lateralized modulation. This discrepancy raises an open

question for future research on the specific role of pre-stimulus (lateralized) alpha in learned suppression.

Notably, we were able to decode the spatial representation of the high-probability location from pre-stimulus alpha oscillations. Furthermore, we successfully differentiated micro-saccades directed toward and away from the high-probability location based on micro-saccade-triggered alpha activity. These findings underscore a potential connection between the human oculomotor system and neural oscillations in the dorsal attention network[51]. We speculate that alpha oscillations might reflect enhanced feedforward control mechanisms over high-probability locations by resetting their phase to facilitate micro-saccade initiation[52], as evidenced by the observed alpha activity preceding micro-saccade-triggered neural oscillations. Future research should investigate how the brain integrates oculomotor and dorsal neural dynamics to enhance attentional processes and optimize behavioral responses in complex environments.

Overall, our findings expand the traditional understanding of micro-saccades by demonstrating their involvement in high-level cognitive processes, particular their role in statistically learned suppression. Within the context of attentional selection, we uncovered potential neural mechanisms reflecting learned suppression, contributing to the ongoing debate on whether attentional processing precedes suppression[12]. Our results suggest that learned suppression is reactive and only occurs after attention is directed to the location that needs to be suppressed. Future research should explore how anticipatory micro-saccades interact with neural activities originating from deep brain structures, particularly the hippocampus, which is recognized as central to the rapid extraction of learned regularities from the environment[53,54].

Methods

Participants

Twenty-eight volunteers (12 males, mean age = 20 years old) participated in the

experiment for monetary compensation (¥80 per hour). One participant was excluded from analysis due to poor EEG recording quality. The sample size was predetermined based on previous studies[6,16]. All participants had normal color vision and normal or corrected-to-normal visual acuity, and provided written informed consent before testing. The study was approved by the ethics review committee of South China Normal University (2020-3-013).

Apparatus and stimuli

Each stimulus contained a vertical or horizontal white line (0.2° × 1.4°, Red-green-blue [RGB]: 255,255,255) inside, presented against a black background (RGB: 0, 0, 0). The enclosing shapes included a circle with a radius of 1° or a diamond subtending 2° × 2°, outlined in either red (RGB: 255, 0, 0) or green (RGB:0, 255, 0) outline. These stimuli were presented on an imaginary circle with a radius of 4°, centered around a white fixation cross (RGB: 255, 255, 255) measuring 1° × 1°. Stimulus presentation and response registration were controlled by custom Python scripts. Participants were seated in a dimly lit laboratory, positioned in front of a 27-in. LCD monitor at a viewing distance of 65 cm.

Procedure and design

We employed the classic additional singleton task[6,35], which required participants to search for a shape singleton while ignoring a color distractor singleton. At the beginning of each trial, a self-paced drift check was performed, following by a central fixation cross remaining visible throughout a trial. After 2500 ms, the search array was presented for 3000 ms or until a response was registered. Participants were instructed to identify the unique shape (target) in the search array, ignoring the salient color singleton (salient distractor), and indicate the orientation (horizontal or vertical) of the line inside the target by pressing the "left" or "up" key, respectively. The experiment contains 10 blocks, with each comprised 120 trials, consisting of 40 distractor-absent trials and 80 distractor-present trials. Participants were required to maintain fixation at the display center throughout each trial. Feedback messages were

presented for trials where participants blinked, failed to respond, pressed a wrong key, or deviated their gaze more than 2° from the fixation cross. Such trials were later retested in a random order after completing the remaining trials.

Due to the predominance of horizontal micro-saccades[29], the salient distractor appeared on either the left (180°) or right (0°) side (counterbalanced across participants) with a probability of 65%, while it appeared at other five locations with an even probability of 7% (see Fig. 1b). The target was presented at each location with equal probability. A brief 10-trial practice session was conducted before the main experiment. Trials on which the response times (RTs) were faster than 200 ms and slower than 2000 ms were excluded from analyses (1.31% of trials).

Eye-tracking acquisition

The eye tracker (EyeLink Portable Duo, SR Research) was positioned approximately 10 cm in front of the monitor, with participants seated ~55 cm away from the screen. Gaze positions (horizontal and vertical) were continuously recorded at a sampling rate of 1000 Hz. Before the beginning of each block, the eye-tracking system was calibrated and validated according to standard protocols provided by the EyeLink software.

Micro-saccade detection

Micro-saccades were detected using an improved version[55] of the algorithm initially proposed by Engbert and Kliegl[29]. Horizontal and vertical eye positions were mapped onto a velocity space, with a relatively low velocity threshold of 3 standard deviations used to detect micro-saccades[55]. For a detected saccade to be considered as a micro-saccade, it needed to exhibit temporal overlap between both eyes, a minimum duration of 5 ms, and an amplitude below 1°. A 100 ms sliding window (stepped in 1-ms intervals) was employed to assess the micro-saccade rate following the fixation onset[29,55,56].

We categorized the micro-saccades based on their directional vectors into 12 distinct bins, with each bin defined by angular ranges of 30° (specifically: 345°-15°,

15°-45°, 45°-75°, 75°-105°, 105°-135°, 135°-165°, 165°-195°, 195°-225°, 225°-255°, 255°-285°, 285°-315°, and 315°-345°). This binning approach allowed for a detailed analysis of the directional distribution of micro-saccades, facilitating the assessment of their role in learned suppression.

Since half participants in the learning condition had high-probability locations on the left side of the visual field while others had them on the right side, we aligned the data by mirroring the directional bins for those participants with high-probability locations on the right side. This mirroring ensured that horizontal micro-saccades within the 345°-15° bin were treated as micro-saccades directed away from the high-probability location, while those within the 165°-195° bin were treated as directed toward the high-probability location.

EEG Recording and Preprocessing

EEG data were recorded using 64 Ag/AgCl active electrodes connected to BrainAmp amplifiers (Brain Products, Munich, Germany), positioned based on the extended 10-20 system, and sampled at 500 Hz. Specifically, electrodes T8 and FT10 were used to record the vertical electro-oculogram (EOG), electrodes T7 and FT9 were used to record the horizontal EOG, electrodes TP9 and TP10 were used to record signals from mastoids, and electrode FCz was used as on-line reference.

The data underwent re-referencing to the mean of the left and right mastoids and were subjected to high-pass filtering with a cutoff frequency of 1.5 Hz (for independent component analysis [ICA] only) and 0.1 Hz (for subsequent analyses). Continuous EEG recordings were epoched from − 3500 to 2500 ms relative to the search array onset. Malfunctioning electrodes were visually detected and temporally removed from the data; and a 110 − 140Hz bandpass filter was used to capture muscle activity and allowed for variable z-score cutoffs per participant based on the within-subject variance of z-scores. Following trial rejection, ICA was conducted solely on the clean electrodes, and components capturing eye blinks, eye movement, or other non-brain-driven artifacts were visually inspected and removed, alongside the

vertical and horizontal EOG signals. Subsequently, we interpolated the malfunctioning electrodes identified earlier.

Time-frequency analysis

Preprocessed EEG data were segmented into event-related epochs (-3500 ms to 2500 ms relative to search array onset; avoiding edge artifacts), and were bandpass filtered using a two-way least-squares finite-impulse-response filter ("eegfilt.m" from EEGLAB Toolbox[57]). A Hilbert transform (MATLAB Signal Processing Toolbox) was applied to the bandpass-filtered data to produce the complex analytic signal, $z(t)$, of the filtered EEG signals, $f(t)$:

$$z(t) = f(t) + i\tilde{f}(t)$$

where $\tilde{f}(t)$ is the Hilbert transform of $f(t)$, and $i = \sqrt{-1}$. The complex analytic signal was extracted for each electrode.

We applied the Hilbert transform on the bandpass filtered data from 8 Hz to 14 Hz in steps of 1 Hz. Total alpha power was computed by squaring the complex magnitude of complex analytic signal, followed by averaging it across trials. The power was then z-scored by subtracting the average value and dividing by the standard deviation across all trials within the focused epoch (-2500 ms to 1000 ms). We selected 17 posterior parietal and occipital electrodes (Pz, P3, P7, O1, Oz, O2, P4, P8, P1, P5, PO7, PO3, POz, PO4, PO8, P6, P2) for further analysis.

Micro-saccade-triggered activity

To analyze micro-saccade-triggered activity within the EEG signals, we first aligned all EEG data to the onset of micro-saccades within the pre-stimulus period (-2500 ms to 0 ms relative to search array onset). The EEG data were then epoched from -1000 ms to 2000 ms relative to the onset of each micro-saccade. We applied the same Hilbert transform methods as before, but in this case, we focused on calculating evoked (phase-locked) power, instead of total power. This allowed us to examine neural responses that were phase-locked to the micro-saccade events and to capture

the specific neural dynamics associated with these tiny eye movements.

Inverted encoding model

To reconstruct spatial representation of distractor locations, we applied an inverted encoding model (IEM) to estimate spatial channel-tuning functions (CTFs) from alpha-band activity over time[37]. The assumption is that the oscillatory activity recorded from each electrode reflects the weighted sum of six spatially selective channels, each tuned to a specific angular position (corresponding to the possible locations of the distractors). The response profile of each spatial channel across distractor locations was modeled as a half sinusoid raised to the seventh power:

$$R = \sin(0.5\theta)^7$$

where θ is the angular location (ranging from 0° to 359°), and R is the response of the spatial channel in arbitrary units. To ensure that the peak response of each spatial channel was centered over one of the six distractor locations, the response profile was circularly shifted.

We partitioned the data randomly into independent sets of training (2/3trials) and test data (1/3 trials) following a cross-validation routine. The training data (B; m electrodes × n locations) were used to estimate the weights that approximated the possible contributions of the six spatial channels to the observed oscillatory activity measured at each electrode and distractor location. We defined C (k channels × n locations) as a matrix of the predicted response of each spatial channel (estimated by the basic function for that channel) for each location; and W (m electrodes × k channels) as a weight matrix characterizing a linear mapping from "channel space" to "electrode space". We further described the relationships between B, C, and W in a linear model as follows:

$$B = WC$$

Using the weight matrix (W) obtained via least-squares estimation, we inverted the model to transform the observed test data into estimated channel responses (i.e.,

reconstructed CTFs). This IEM routine was iterated 10 times to minimize the influence of idiosyncrasies specific to any trial assignment and obtain averaged channel-response profiles. We then calculated the slope of the CTFs via linear regression to quantify attentional processing for different distractor locations over time. Higher slope values indicate greater spatial selectivity while lower values indicate less spatial selectivity. Distractor-tuning CTFs were reconstructed according to the distractor location across participants. The data were down-sampled to 50 Hz to reduce computational time.

Classification analysis

We employed a Support Vector Machine with error-correcting output codes (SVM-ECOC) to create two classifiers[58], aimed at classifying whether anticipatory micro-saccades were directed toward or away from the high-probability location, based on micro-saccade-triggered activities. This was implemented using the MATLAB function 'fitcecoc'. Due to significant variability in the number of micro-saccade-triggered EEG epochs across participants, subject-level analysis was not feasible. However, assuming that responses from the same electrodes across different participants should be consistent under the same cognitive process, we concatenated all epochs across all participants. Specifically, with a threefold cross-validation procedure, 2/3 of epochs were randomly selected for model training, and the remaining 1/3 were used for testing. To reduce biases from specific grouping, 100 iterations of this procedure were conducted with distinct randomizations. Additionally, the data were down-sampled to 50 Hz to reduce computational time.

Statistical analysis

To correct multiple comparisons, we used a cluster-based permutation test against a null distribution shuffled from 1000 iterations (following the Monte Carlo randomization procedure) for each temporal-level analysis. Specifically, one-sample t-tests were performed across participants for conditional difference (or single-condition activity) against zeros to identify above-chance activity. Time

windows with t values larger than a threshold (p = .05) were combined into contiguous clusters based on adjacency. The cluster statistics was defined as the sum of the t values within each cluster. Cluster statistics were defined as the sum of the t values within each cluster. The null distribution was formed by randomly permuting condition labels for 1000 times to get the largest clusters per iteration. Clusters were determined to be significant if the cluster statistics were larger than the 95th percentile of the null distribution.

For the classification analysis, since subject-level t-tests could not be performed, we employed a non-parametric approach by repeating the SVM-ECOC procedure 1000 times while randomizing the labels of micro-saccade directions. For each time point, a null distribution of classification accuracy was generated. Classification accuracies were determined to be significant if they exceeded the 95th percentile of the null distribution.

References

1. Jiang, Y. V., Swallow, K. M., Rosenbaum, G. M. & Herzig, C. Rapid acquisition but slow extinction of an attentional bias in space. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 87–99 (2013).
2. Sha, L. Z., Remington, R. W. & Jiang, Y. V. Short-term and long-term attentional biases to frequently encountered target features. *Atten. Percept. Psychophys.* **79**, 1311–1322 (2017).
3. Ferrante, O. *et al.* Altering spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex* **102**, 67–95 (2018).
4. Theeuwes, J., Bogaerts, L. & Moorselaar, D. van. What to expect where and when: how statistical learning drives visual selection. *Trends Cogn. Sci.* **26**, 860–872 (2022).
5. Theeuwes, J. Attentional Capture and Control. *Annu. Rev. Psychol.* **76**, 251–273 (2025).
6. Wang, B. & Theeuwes, J. Statistical regularities modulate attentional capture. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 13–17 (2018).
7. Wang, B. & Theeuwes, J. How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Atten. Percept. Psychophys.* **80**, 860–870 (2018).
8. Wang, B. & Theeuwes, J. Statistical regularities modulate attentional capture independent of search strategy. *Atten. Percept. Psychophys.* **80**, 1763–1774 (2018).
9. Goschy, H., Bakos, S., Müller, H. J. & Zehetleitner, M. Probability cueing of distractor locations: both intertrial facilitation and statistical learning mediate interference reduction. *Front. Psychol.* **5**, (2014).
10. Leber, A. B., Gwinn, R. E., Hong, Y. & O'Toole, R. J. Implicitly learned suppression of irrelevant spatial locations. *Psychon. Bull. Rev.* **23**, 1873–1881 (2016).
11. Sauter, M., Liesefeld, H. R., Zehetleitner, M. & Müller, H. J. Region-based shielding of visual search from salient distractors: Target detection is impaired with same- but not different-dimension distractors. *Atten. Percept. Psychophys.* **80**, 622–642 (2018).
12. Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W. & Theeuwes, J. Progress toward resolving the attentional capture debate. *Vis. Cogn.* **29**, 1–21 (2021).
13. Huang, C., Vilotijević, A., Theeuwes, J. & Donk, M. Proactive distractor suppression elicited by statistical regularities in visual search. *Psychon. Bull. Rev.* **28**, 918–927 (2021).
14. Huang, C., Donk, M. & Theeuwes, J. Attentional suppression is in place before display onset. *Atten. Percept. Psychophys.* **85**, 1012–1020 (2023).
15. Kong, S., Li, X., Wang, B. & Theeuwes, J. Proactively location-based suppression elicited by statistical learning. *PLOS ONE* **15**, e0233544 (2020).
16. Wang, B., van Driel, J., Ort, E. & Theeuwes, J. Anticipatory Distractor Suppression Elicited by Statistical Regularities in Visual Search. *J. Cogn.*

Neurosci. **31**, 1535–1548 (2019).

17. Munneke, J., Heslenfeld, D. J., Usrey, W. M., Theeuwes, J. & Mangun, G. R. Preparatory Effects of Distractor Suppression: Evidence from Visual Cortex. *PLOS ONE* **6**, e27700 (2011).
18. Richter, D., Moorselaar, D. van & Theeuwes, J. Proactive distractor suppression in early visual cortex. 2024.04.03.587747 Preprint at https://doi.org/10.1101/2024.04.03.587747 (2024).
19. Beck, V. M., Luck, S. J. & Hollingworth, A. Whatever you do, don't look at the . . .: Evaluating guidance by an exclusionary attentional template. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 645–662 (2018).
20. Humphreys, G. W., Stalmann, B. J. & Olivers, C. An analysis of the time course of attention in preview search. *Percept. Psychophys.* **66**, 713–730 (2004).
21. Lahav, A., Makovski, T. & Tsal, Y. White bear everywhere: Exploring the boundaries of the attentional white bear phenomenon. *Atten. Percept. Psychophys.* **74**, 661–673 (2012).
22. Lahav, A. & Tsal, Y. Allocating Attention to Distractor Locations is Based on Top-Down Expectations. *Q. J. Exp. Psychol.* **66**, 1873–1880 (2013).
23. Makovski, T. APA PsycNet Buy Page. https://psycnet.apa.org/buy/2018-52923-001 (2019).
24. Moher, J. & Egeth, H. E. The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Atten. Percept. Psychophys.* **74**, 1590–1605 (2012).
25. Tsal, Y. & Makovski, T. The attentional white bear phenomenon: The mandatory allocation of attention to expected distractor locations. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 351–363 (2006).
26. Won, B.-Y., Kosoyan, M. & Geng, J. J. Evidence for second-order singleton suppression based on probabilistic expectations. *J. Exp. Psychol. Hum. Percept. Perform.* **45**, 125–138 (2019).
27. Klein, R. M. Inhibition of return. *Trends Cogn. Sci.* **4**, 138–147 (2000).
28. Wang, B. & Theeuwes, J. Salient distractor processing: inhibition following attentional capture. *Trends Cogn. Sci.* **28**, 593–594 (2024).
29. Engbert, R. & Kliegl, R. Microsaccades uncover the orientation of covert attention. *Vision Res.* **43**, 1035–1045 (2003).
30. Yuval-Greenberg, S., Merriam, E. P. & Heeger, D. J. Spontaneous Microsaccades Reflect Shifts in Covert Attention. *J. Neurosci.* **34**, 13693–13700 (2014).
31. Jiang, Y. V., Won, B.-Y. & Swallow, K. M. First saccadic eye movement reveals persistent attentional guidance by implicit learning. *J. Exp. Psychol. Hum. Percept. Perform.* **40**, 1161–1173 (2014).
32. van Ede, F., Chekroud, S. R. & Nobre, A. C. Human gaze tracks attentional focusing in memorized visual space. *Nat. Hum. Behav.* **3**, 462–470 (2019).
33. Denison, R. N., Yuval-Greenberg, S. & Carrasco, M. Directing Voluntary Temporal Attention Increases Fixational Stability. *J. Neurosci.* **39**, 353–363 (2019).
34. De Vries, E. & Van Ede, F. Microsaccades Track Location-Based Object

Rehearsal in Visual Working Memory. *eneuro* **11**, ENEURO.0276-23.2023 (2024).

35. Theeuwes, J. Perceptual selectivity for color and form. *Percept. Psychophys.* **51**, 599–606 (1992).
36. Lin, R. *et al.* Neural evidence for attentional capture by salient distractors. *Nat. Hum. Behav.* **8**, 932–944 (2024).
37. Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K. & Awh, E. The topography of alpha-band activity tracks the content of spatial working memory. *J. Neurophysiol.* **115**, 168–177 (2016).
38. Dimigen, O., Valsecchi, M., Sommer, W. & Kliegl, R. Human Microsaccade-Related Visual Brain Responses. *J. Neurosci.* **29**, 12321–12331 (2009).
39. Meyberg, S., Werkle-Bergner, M., Sommer, W. & Dimigen, O. Microsaccade-related brain potentials signal the focus of visuospatial attention. *NeuroImage* **104**, 79–88 (2015).
40. Liu, B., Nobre, A. C. & van Ede, F. Microsaccades transiently lateralise EEG alpha activity. *Prog. Neurobiol.* **224**, 102433 (2023).
41. Chen, L., Cichy, R. M. & Kaiser, D. Coherent categorical information triggers integration-related alpha dynamics. *J. Neurophysiol.* **131**, 619–625 (2024).
42. Abeles, D., Amit, R., Tal-Perry, N., Carrasco, M. & Yuval-Greenberg, S. Oculomotor inhibition precedes temporally expected auditory targets. *Nat. Commun.* **11**, 3524 (2020).
43. Badde, S., Myers, C. F., Yuval-Greenberg, S. & Carrasco, M. Oculomotor freezing reflects tactile temporal expectation and aids tactile perception. *Nat. Commun.* **11**, 3341 (2020).
44. Dankner, Y., Shalev, L., Carrasco, M. & Yuval-Greenberg, S. Prestimulus Inhibition of Saccades in Adults With and Without Attention-Deficit/Hyperactivity Disorder as an Index of Temporal Expectations. *Psychol. Sci.* **28**, 835–850 (2017).
45. Hung, S.-C. & Carrasco, M. Microsaccades as a long-term oculomotor correlate in visual perceptual learning. *Psychon. Bull. Rev.* **30**, 235–249 (2023).
46. Serences, J. T., Schwarzbach, J., Courtney, S. M., Golay, X. & Yantis, S. Control of Object-based Attention in Human Cortex. *Cereb. Cortex* **14**, 1346–1357 (2004).
47. Ruff, C. C. & Driver, J. Attentional Preparation for a Lateralized Visual Distractor: Behavioral and fMRI Evidence. *J. Cogn. Neurosci.* **18**, 522–538 (2006).
48. Theeuwes, J., Olivers, C. N. L. & Chizk, C. L. Remembering a Location Makes the Eyes Curve Away. *Psychol. Sci.* **16**, 196–199 (2005).
49. Van der Stigchel, S., Meeter, M. & Theeuwes, J. The spatial coding of the inhibition evoked by distractors. *Vision Res.* **47**, 210–218 (2007).
50. Jensen, O. & Mazaheri, A. Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Front. Hum. Neurosci.* **4**, (2010).
51. Popov, T., Gips, B., Kastner, S. & Jensen, O. Spatial specificity of alpha oscillations in the human visual system. *Hum. Brain Mapp.* **40**, 4432–4440

(2019).

52. Barczak, A. *et al.* Dynamic Modulation of Cortical Excitability during Visual Active Sensing. *Cell Rep.* **27**, 3447-3459. (2019).
53. Chun, M. M. & Phelps, E. A. Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nat. Neurosci.* **2**, 844–847 (1999).
54. Turk-Browne, N. B., Scholl, B. J., Chun, M. M. & Johnson, M. K. Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness. *J. Cogn. Neurosci.* **21**, 1934–1945 (2009).
55. Engbert, R. & Mergenthaler, K. Microsaccades are triggered by low retinal image slip. *Proc. Natl. Acad. Sci.* **103**, 7192–7197 (2006).
56. Laubrock, J., Engbert, R. & Kliegl, R. Microsaccade dynamics during covert attention. *Vision Res.* **45**, 721–730 (2005).
57. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
58. Bae, G.-Y. & Luck, S. J. Dissociable Decoding of Spatial Attention and Working Memory from EEG Oscillations and Sustained Potentials. *J. Neurosci.* **38**, 409–422 (2018).