Presearch Attentional Focus Supports Learned Suppression in Visual Search

Yinfei Zhou[1,2,3], Lishuang Wang[2,3], Xinyu Li[1], Jan Theeuwes[4], and Benchi Wang[2,3]

Abstract

■ Salient distractors often capture our attention, disrupting ongoing tasks. Recent studies suggest that, through statistical learning, prior experiences regarding distractor locations can reduce distraction by suppressing their corresponding locations. However, the proactive neural mechanisms supporting this learned suppression remain unclear. Our findings demonstrate that participants learn to suppress locations that are more likely to contain distractors relative to other locations. Using frequency tagging in EEG recordings, we observed significantly different tagging responses between high- and low-probability locations, along with a general decrease in alpha power (8–12 Hz) before search onset. Notably, the higher tagging frequency power at high-probability locations suggests that participants allocated greater attentional focus to these locations in anticipation of the search. These results suggest that anticipatory attentional deployment precedes the suppression of high-probability distractor locations after the onset of visual search. ■

INTRODUCTION

In everyday life, we embrace considerable visual information from the environment but cannot process all of it at once. To act and behave in a goal-directed manner, we focus our limited resources on relevant information and filter out distracting information (Johnston & Dark, 1986). Although selective attention is an effective mechanism that determines what we see and act upon, our attention may be captured by salient distractors. One efficient way to suppress these distractors is learning from past experiences (regularities) regarding distractors (Theeuwes, Bogaerts, & van Moorselaar, 2022; Wang & Theeuwes, 2018a, 2018b) via a mechanism called statistical learning (SL; Turk-Browne, Junge, & Scholl, 2005). For example, behavioral studies have established that learning to expect where distractors are most probable during visual search can help suppress these distractors and facilitate upcoming search (Ferrante et al., 2018; Sauter, Liesefeld, Zehetleitner, & Müller, 2018; Wang & Theeuwes, 2018a, 2018b, 2018c; Leber, Gwinn, Hong, & O'Toole, 2016; Goschy, Bakos, Müller, & Zehetleitner, 2014). Although this effect has been established many times, the neural mechanism regarding learned suppression remains less clear, especially whether this type of suppression can be implemented proactively, by modulating different distractor locations separately.

To this aim, we adopted EEG recordings in humans when performing the additional singleton task (Theeuwes, 1992), in which participants were required to search for a unique shape (target) while ignoring a salient distractor. Notably, the salient distractor was presented more often in one specific location than in other three locations in the learning condition, introducing SL that should result in the suppression of the frequent distractor location. To track the prestimulus attentional distribution due to learned suppression, a technique called frequency tagging (Zhigalov, Herring, Herpers, Bergmann, & Jensen, 2019; Malinowski, Fuchs, & Müller, 2007; Ding, Sperling, & Srinivasan, 2006; Müller, Malinowski, Gruber, & Hillyard, 2003; Morgan, Hansen, & Hillyard, 1996) was used to separately assess frequency tagging response to different distractor locations. If learned suppression operates through proactive attention mechanism, we expect corresponding neural activity would emerge before the stimulus onset. We also measured alpha activity simultaneously, as alpha-band activity is considered to be a neural signature of functional inhibition (Foxe & Snyder, 2011; Jensen & Mazaheri, 2010; Klimesch, Sauseng, & Hanslmayr, 2007) and is strongly modulated by spatial attention (Sauseng et al., 2005).

METHODS

Participants

Twenty-eight undergraduate students (20 women and eight men with a mean age of 20 years, predetermined based on Wang, van Driel, Ort, & Theeuwes, 2019) participated in this experiment. The study was approved by the ethics committee at South China Normal University (2020-

[1]Zhejiang Normal University, [2]Key Laboratory of Brain, Cognition and Education Sciences, [3]South China Normal University, [4]Vrije Universiteit Amsterdam

© 2025 Massachusetts Institute of Technology

Journal of Cognitive Neuroscience 37:12, pp. 2493–2502
https://doi.org/10.1162/JOCN.a.61

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn.a.61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

3-013), and all participants provided written informed consent before taking part. They were all right-handed, had normal or corrected-to-normal vision, and were financially compensated (¥70 per hour).

Apparatus and Stimuli

Stimulus presentation and behavioral data collection were controlled by custom scripts written in Python and run on a 27-in. computer with a 1920 × 1080 resolution and a refresh rate of 60 Hz. Participants were seated in a sound-attenuated and dimly lit laboratory at a viewing distance of 60 cm. All stimuli were presented against a black background (red–green–blue [RGB]: 0, 0, 0) with a white fixation dot (RGB: 255, 255, 255; radius: 0.25°) presented in the center of the screen. Four gray (RGB: 128, 128, 128) placeholders (2° × 2°) were presented 4° away from the fixation. The search array consisted of four discrete stimuli with different shapes (one circle among three unfilled diamonds or vice versa) superimposed on the gray placeholders. The circle had a radius of 1°, and the diamond was subtended 2° × 2°. Each shape had a red (RGB: 255, 0, 0) or green (RGB: 0, 255, 0) outline and a vertical or horizontal white line (0.2° × 1.4°) inside (see Figure 1A).

Procedure and Design

During the experiment, participants were instructed to search for a unique shape (a circle among diamonds or vice versa) while ignoring a salient distractor (red or green, which was noticeably different in color from the other search elements). Participants had to indicate the orientation (horizontal or vertical) of the line within the target shape. They were asked to respond as fast and accurately as possible. Each trial began with a fixation dot (100–300 msec), followed by four gray placeholders that flickered (tagged) for 1250 msec. Subsequently, the search array was presented for 2000 msec, with those four gray placeholders that continued to flicker. Notably, participants were explicitly instructed to ignore the tagging placeholders, ensuring their attention remained focused on the primary visual search task and not diverted to the tagging backgrounds. The intertrial interval was between 500 and 700 msec (see Figure 1A).

Before the experiment, participants received verbal instructions and completed 15 practice trials to be familiar with placeholder tagging and the experimental task. Each participant completed five blocks for the baseline condition and 10 blocks for the learning condition. Each block was composed of 120 trials, with 40 distractor-absent trials

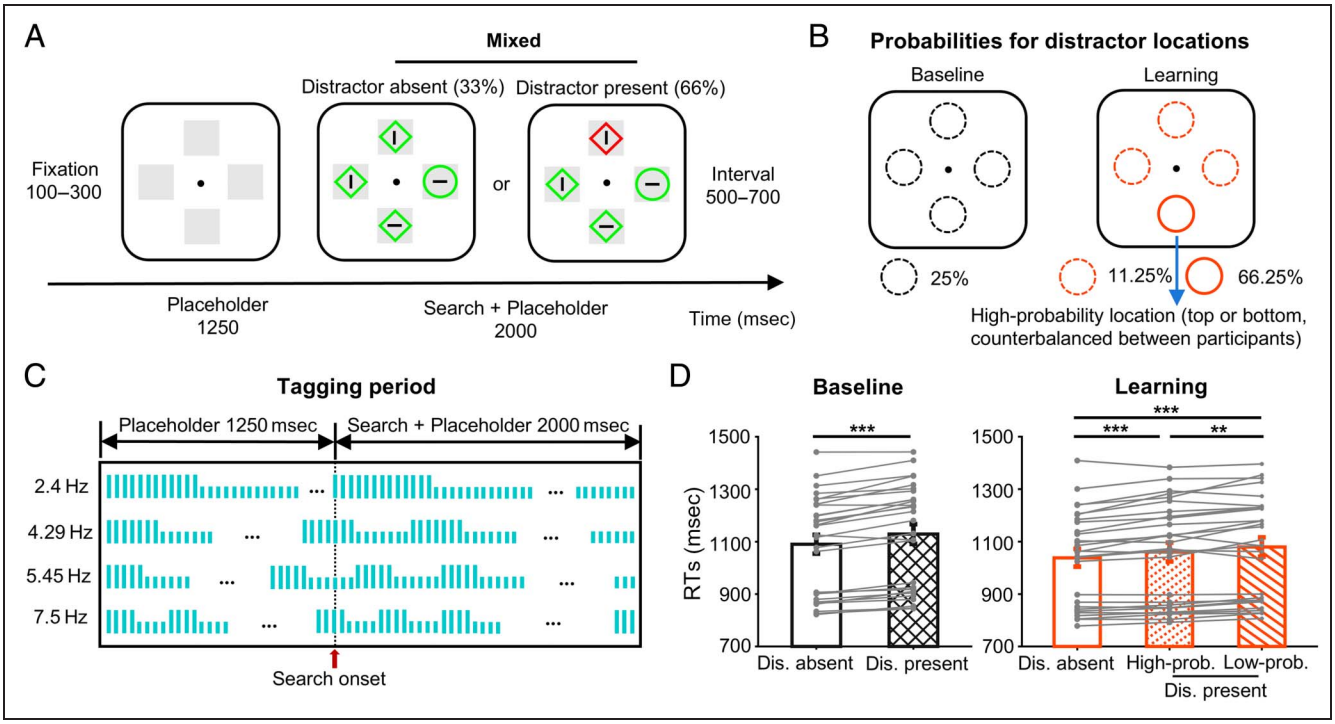


Figure 1. (A) Experimental procedure. Each trial started with a fixation dot followed by four tagging placeholders for 1250 msec. Then, the search array was presented with all tagging placeholders simultaneously for 2000 msec at different frequencies. Participants were asked to find the unique shape (in this case, the circle to the right) and report the orientation of the line inside (in this case, horizontal). Distractor-present and distractor-absent trials were mixed within each block. (B) Probabilities for distractor locations. In the baseline condition, the salient distractor was equally presented in four locations, with a probability of 25% each. In the learning condition, the salient distractor was presented in either the top or bottom location (which was counterbalanced between participants) with a probability of 66.25%, whereas it was equally presented in other three locations with a probability of 11.25% each. (C) Temporal sequence for different tagging placeholders at different frequencies, that is, 2.4, 4.29, 5.45, and 7.5 Hz (counterbalanced between participants). Each mini vertical bar represents a frame (16.7 msec), with the longer and short ones representing the onset and offset of the placeholders respectively. (D) Mean RTs in the baseline and learning conditions. The gray solid dots represent individual RTs. Error bars show ±1 *SEM*. Dis. = distractor; prob. = probability. $^{**}p < .01$, $^{***}p < .001$.

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn.a.61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

and 80 distractor-present trials. In the baseline condition, which was performed before the learning condition to avoid any contamination from learning, the salient distractor was equally presented in four locations, with a 25% probability each. In the learning condition, the salient distractor was presented in either the top or bottom location (which was counterbalanced between participants) with a probability of 66.25%, whereas it was equally presented in other three locations with a probability of 11.25% each. The target was randomly selected from nondistractor locations in both conditions, resulting in no clue for participants to figure out where the target would be. More blocks for the learning condition were designed to examine whether any effect observed was due to that the baseline condition was tested first for all participants (see the last paragraph in the Results section for more details).

Notably, we deliberately chose not to present the high-probability location laterally because alpha lateralization effects could be confounded by bilateral electrode changes, rather than specifically reflecting modulation of the high-probability location. For instance, findings from Wang et al. (2019) suggest that any observed increase in alpha power contralateral to the high-probability location might partially arise from decreases in power contralateral to the low-probability location.

Frequency Tagging

Frequency tagging refers to the presentation of visual stimuli that tag at specific frequencies, resulting in steady-state visual evoked fields that can be observed in EEG signals (Norcia, Appelbaum, Ales, Cottereau, & Rossion, 2015). Here, we employed four different frequencies (2.4, 4.29, 5.45, and 7.5 Hz) for each placeholder before and after the search display onset, allowing us to track attentional distribution over time. These frequencies were counterbalanced between participants. The criteria for selecting these frequencies were as follows: (1) to avoid any overlap with alpha oscillation (8–12 Hz), which was one of the primary foci in our study; (2) to prevent any high frequencies (>12 Hz) that may cause discomfort for participants; (3) to ensure that each frequency was not a harmonic of another; and (4) to ensure that the number of frames corresponding to the frequency is an integer. Our monitor's refresh rate was 60 Hz, resulting in 25, 14, 11, and 8 frames for 2.4, 4.29, 5.45, and 7.5 Hz, respectively (see Figure 1B). These frequencies corresponding to different locations were counterbalanced between participants.

EEG Recording and Preprocessing

EEG data were acquired using 64 Ag/AgCl active electrodes connected to BrainAmp amplifiers (Brain Products), placed according to the extended 10–20 system and digitalized at 500 Hz. Here, electrodes T8 and FT10 were used to record vertical EOG, and electrodes T7 and FT9 were used to record horizontal EOG; moreover, electrodes TP9 and TP10 were used to record signals from mastoids, and electrode Fz was used as online reference.

The data were rereferenced to the mean of left and right mastoids and high-pass filtered using a cutoff of 1.5 Hz (for independent component analysis [ICA] only) and 0.1 Hz (for final analyses). Continuous EEG was epoched from −2000 to 4750 msec relative to the placeholder onset. Malfunctioning electrodes were visually detected and temporally removed from the data, and a 110- to 140-Hz 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. After trial rejection, ICA was performed on the clean electrodes only. Together with the vertical and horizontal EOG signals, we visually inspected and removed ICA components that captured eye blinks, eye movement, or other artifacts that were clearly not brain-driven activity. Afterward, we interpolated the malfunctioning electrodes identified earlier. Offline preprocessing was performed using MATLAB v2016 (MathWorks Inc.), incorporating EEGLab2020_0 and custom MATLAB scripts.

Time–Frequency Analysis

Preprocessed EEG data were broken into event-related epochs (−2000 to 4750 msec relative to placeholder onset; avoiding edge artifacts from wavelet convolution) and then were convolved with a set of Morlet wavelets with frequencies including 2.4, 4.29, 5.45, 7.5, 8, 9, 10, 11, 12, 20, 25, and 30 Hz. The number of wavelet cycles was logarithmically spaced between 3 and 12 to have a good trade-off between temporal and frequency precision. This might result in a potential overlap between neighboring frequencies. However, the selected frequencies corresponding to different locations were counterbalanced between participants, preventing such potential overlapping effects. Because each frequency selected for the high-probability location was just one of several frequencies selected within the alpha band, it is unlikely that its activity alone can represent the full alpha-band activity. Furthermore, even when we varied the number of cycles (1.5, 1.6, 1.7, 1.8, 1.9, and 2 times the corresponding frequencies) to better differentiate between tagging frequencies and to keep the temporal resolution fixed, the results remained consistent (see Figure S3 in the Supplemental Material for details). The power within our focused time window (0–3250 msec) was then averaged across trials and decibel-transformed relative to a prestimulus period of −500 to −300 msec relative to the placeholder onset. We did not calculate phase-locked activity for further analysis because our task required participants to switch their attention across different search elements to search for the target. As a result, the neural activity produced by SL during visual search was not phase-locked.

On the basis of previous studies, we selected PO3/4, PO7/8, and O1/2 for further analysis (Antonov, Chakravarthi, & Andersen, 2020; Wang et al., 2019). We were specifically

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn_a_61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

interested in the tagging frequencies (2.4, 4.29, 5.45, and 7.5 Hz), alpha frequencies (8, 9, 10, 11, and 12 Hz), three control frequencies (20, 25, and 30 Hz; which was far from the tagging frequencies and alpha oscillation, avoiding potential impact from these key frequencies), and their comparisons.

Cluster-based Permutation Test

To correct multiple comparisons, we used a cluster-based permutation test against a null distribution shuffled from 1000 iterations (following Monte Carlo randomization procedure). Specifically, one-sample t tests were performed across participants for conditional differences against zero to identify above-chance activity, and 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. 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 was larger than the 95th percentile of the null distribution.

RESULTS

Participants performed a visual search task, in which they were instructed to search for a unique shape (target) while ignoring a salient distractor (see Figure 1A). In the baseline condition, the salient distractor was equally presented at all four locations. In the learning condition, the salient distractor was presented more often in either the top or bottom location (which was counterbalanced between participants; high-probability location), and was equally presented in other three locations (low-probability location), producing learned suppression on salient distractors. Notably, four gray placeholders were tagged at different frequencies before and after the search onset, and participants were explicitly told to ignore the tagging background (see Methods for details).

Behavioral Results

Incorrect trials, as well as trials on which the RTs were faster than 200 msec and slower than 2000 msec, or no response was given within 2000 msec, were excluded from analyses (1.29%). Mean RTs for different conditions are presented in Figure 1D. In the baseline condition, compared with distractor-absent trials (1090 msec), paired t tests revealed that the mean RTs were significantly larger for distractor-present trials (1129 msec), $t(27) = 7.55$, $p < .001$, $d = 1.43$, indicating capture by salient distractors. The results on error rates mimicked those for RTs. The error rates were higher for distractor-present trials (5.1%) than distractor-absent trials (3.8%), $t(27) = 3.92$, $p < .001$, $d = 0.74$.

In the learning condition, with Distractor condition (low-probability location, high-probability location, and distractor-absent) as a factor, a repeated-measures ANOVA on mean RTs showed a main effect, $F(2, 54) = 26.88$, $p < .001$, partial $\eta^2 = .5$. Subsequent planned comparisons revealed larger mean RTs for a distractor presented in the high-probability location (1060 msec), $t(27) = 4.76$, $p < .001$, $d = 0.9$, and the low-probability location (1079 msec), $t(27) = 6.4$, $p < .001$, $d = 1.21$, compared to the distractor-absent condition (1038 msec). Crucially, there was a reliable difference between the high- and low-probability locations, $t(27) = 3.37$, $p = .002$, $d = 0.64$, suggesting that salient distractor was suppressed after learning its high-probability location. No significant effect for error rates was observed, $F(2, 54) = 1.22$, $p = .303$, partial $\eta^2 = .04$.

Overall, the results demonstrated the classic attentional capture effect by salient distractors (Lin et al., 2024; Theeuwes, 1992) and further indicated that this capture can be suppressed through the learning of statistical regularities regarding distractor locations (Wang & Theeuwes, 2018a, 2018b, 2018c). A recent study by Kerzel, Balbiani, Rosa, and Huynh Cong (2022) suggested that reduced capture is not solely due to suppression of the high-probability location but also results from increased interference from distractors at low-probability locations (i.e., distractor rarity). In this study, larger capture effects were observed at low-probability locations compared to equal-probability locations, which correspond to the baseline condition in the present study. However, in our study, the capture effect was similar at low-probability locations (41 msec) and in the baseline condition (39 msec). This discrepancy may stem from differences in the experimental design. Unlike the between-subject design used by Kerzel et al. (2022), our within-subject design introduced the baseline condition before the learning condition, allowing participants to practice extensively before the learning condition. This practice likely reduced capture effects, particularly at low-probability locations. Nevertheless, this discrepancy does not undermine our conclusion that the high-probability location was suppressed, a finding consistent with Kerzel et al. (2022).

Notably, compared to previous studies that only four items had on display, we observed much longer RTs (e.g., Wang & Theeuwes, 2020). One possible explanation is that the tagging frames may have reduced the visibility of the shapes, making the task more challenging and contributing to the slower RTs observed.

Frequency Tagging and Alpha Oscillation

In Figure 2A (left), we depict the power spectra across frequencies averaged over the tagging period (from 0 to 3250 msec) for the baseline and learning conditions, as well as their differences. To validate the efficacy of frequency tagging and assess the role of alpha oscillation in the present study, we computed the average power for tagging

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn_a_61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

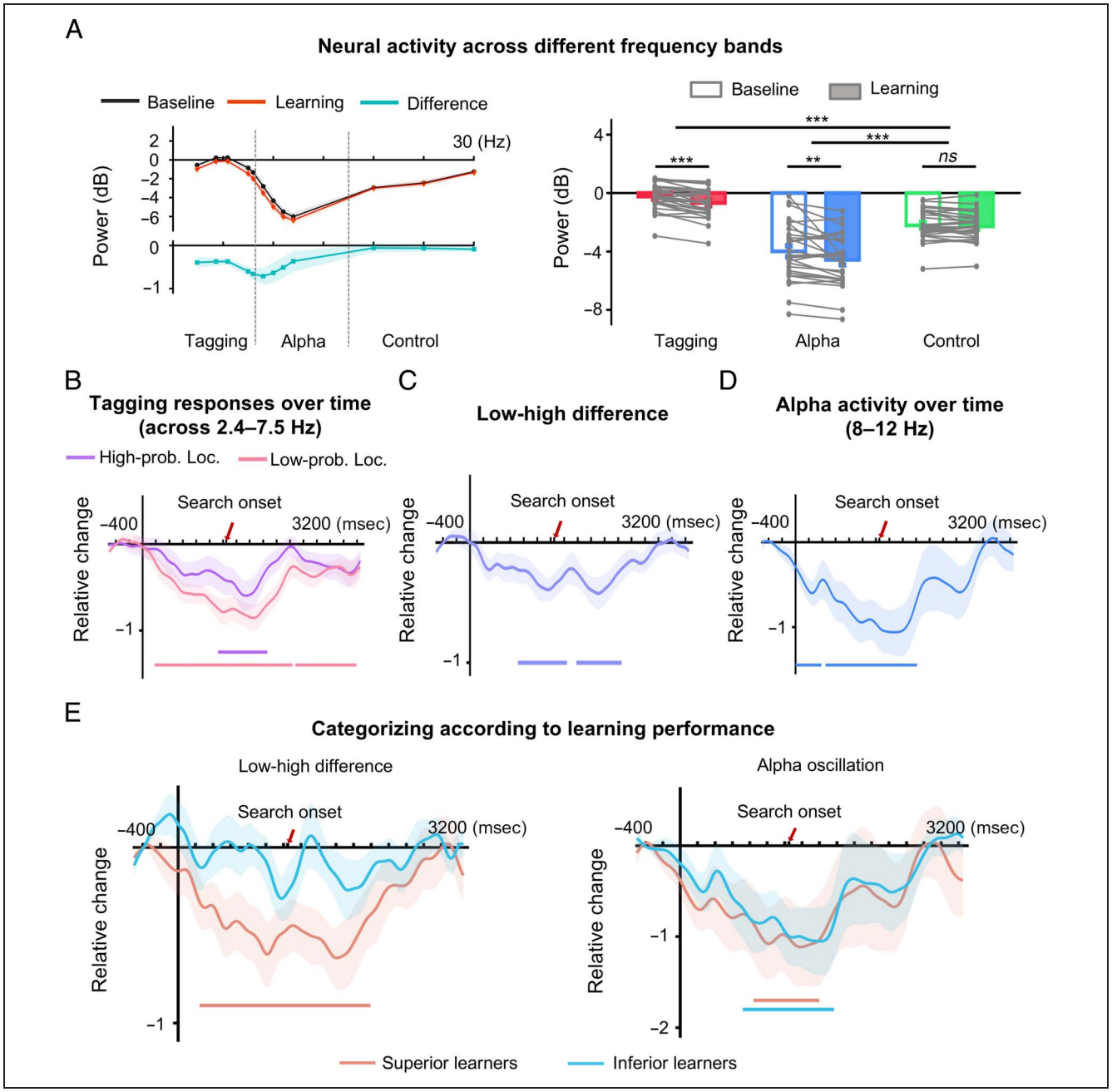


Figure 2. (A) The left graph shows the power spectra across different frequencies averaged over tagging period (from 0 to 3250 msec) for the baseline and learning conditions, and their difference, including tagging frequencies (2.4, 4.29, 5.45, and 7.5 Hz), alpha frequencies (8, 9, 10, 11, and 12 Hz), and control frequencies (20, 25, and 30 Hz). The right graph shows the averaged power for tagging, alpha, and control frequencies in the baseline and learning conditions. The gray solid dots represent individual data. Error bars show ±1 *SEM*. $^{**}p < .01$, $^{***}p < .001$. (B) Time-varied changes (i.e., power difference obtained by subtracting the power observed in the baseline condition from that in the learning condition) for tagging responses toward high- and low-probability distractor locations. (C) The low–high difference (obtaining by subtracting tagging responses toward high-probability location from those toward low-probability locations) indicates the ability to discern the high-probability location from other low-probability locations. (D) Time-varied changes for alpha activity. (E) The low–high difference (left) and alpha oscillation (right) for superior and inferior learners. Colored shadow areas represent ±1 *SEM*, and the horizontal lines represent significant clusters after cluster-based permutation test at $p < .05$. Prob. = probability; Loc. = location.

frequencies (2.4, 4.29, 5.45, and 7.5 Hz), alpha frequencies (8, 9, 10, 11, and 12 Hz), and three control frequencies (20, 25, and 30 Hz). Notably, we did not employ tagging for the search elements; instead, the placeholders (i.e., the stimulus background) were tagged continuously throughout the trial, with participants explicitly instructed to ignore them. The background (i.e., tagging placeholders) served as a distraction for the search task. Previous studies have shown that enhancing attention to a stimulus causes the suppression of unattended (distracting) stimulus simultaneously (Andersen & Müller, 2010; Müller & Hübner, 2002), thereby leading to reduction in frequency power when

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn.a.61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

tagging the unattended stimulus. Thus, we hypothesized that the power for tagging frequencies would be lower than zero, as the tagging placeholders were disregarded throughout the trial. Consistent with this hypothesis, the modulation of the tagging frequencies (averaged across frequencies) was significantly lower than zero, $t(27) = 2.64$, $p = .014$, $d = 0.5$.

Subsequently, with Frequency Type (tagging, alpha, and control frequencies) and Condition (baseline vs. learning) as factors, a repeated-measures ANOVA revealed significant main effects for Frequency Type, $F(2, 54) = 90$, $p < .001$, partial $\eta^2 = .77$, and Condition, $F(1, 27) = 15.63$, $p < .001$, partial $\eta^2 = .37$, along with a noteworthy interaction, $F(2, 54) = 6.11$, $p = .004$, partial $\eta^2 = .19$ (see Figure 1A, right). Further planned comparisons revealed that, in both baseline and learning conditions, suppression for tagging frequencies was weaker than that for control frequencies, both ps $< .001$, indicating tagging responses toward different locations after tagging placeholders. In addition, alpha power exhibited a decrease relative to control frequencies, both ps $< .001$, signifying the involvement of alpha suppression in the present study.

Importantly, distinctions between baseline and learning conditions were evident for tagging frequencies, $t(27) = 5.84$, $p < .001$, $d = 1.1$, and alpha oscillation, $t(27) = 2.99$, $p = .006$, $d = 0.57$, but not for control frequencies, $t(27) = 1.05$, $p = .302$, $d = 0.2$. This shows that both frequency tagging responses and alpha oscillation decrease after the distractor location can be anticipated compared to the baseline condition.

Tracking High- vs. Low-Probability Locations Over Time

To explore the related neural activities associated with learned suppression, we initially calculated the power difference at each time point inside the tagging time window (0–3250 msec). This involved subtracting the power observed in the baseline condition from that in the learning condition, obtaining a pure learning effect characterized by relative changes in power. We categorized these power differences into three components: tagging responses tracking the high-probability location, tagging responses tracking other low-probability locations, and alpha oscillation separately. As depicted in Figure 2B, as time progressed, tagging responses occurred in the intervals of 180–2238 and 2282–3192 msec for low-probability locations, whereas they occurred in the interval of 1126–1862 msec for the high-probability location (cluster-based permutation test, $p < .05$; see Figure S1 in the Supplemental Material for the baseline condition), primarily falling within the presearch time window (0–1250 msec). Notably, the difference between tagging responses toward high- and low-probability locations (referred to as the low–high difference, calculated by subtracting tagging responses toward those high-probability locations from that toward low-probability locations), occurred in the intervals of 714–1442 and 1580–2252 msec (cluster-based permutation test, $p < .05$; see Figure 2C). This indicates that participants can discern the high-probability location from other low-probability locations even before the commencement of the search, by showing a stronger frequency power in tagging response toward the high-probability location. Moreover, alpha activity was reduced in the learning condition in the intervals of 0–386 and 448–1812 msec relative to the placeholder onset (cluster-based permutation test, $p < .05$; see Figure 2D). These results suggest proactive attentional mechanisms modulating high- and low-probability locations separately.

In addition, this discrimination in tagging responses between high- and low-probability locations aligns with participants' behavioral responses. Utilizing a median split based on behavioral indices of learning (i.e., mean RTs for presenting a distractor in low-probability locations minus those in the high-probability location), we categorized participants into "superior learners" and "inferior learners" ($n = 14$ for each group). The results unveiled that the significant low–high difference persisted for superior learners in the interval of 246–2192 msec, whereas it became weak and did not reach significance for inferior learners (see Figure 2E, left). However, the presearch alpha activity existed for both superior learners (842–1598 msec) and inferior learners (720–1766 msec), with no significant difference detected between the two groups (cluster-based permutation test, $p < .05$; see Figure 2E, right). These findings underscore the role of tagging responses in distinguishing high- and low-probability locations within learned suppression. Yet, the role of alpha oscillation remains ambiguous. We will return to this discussion later.

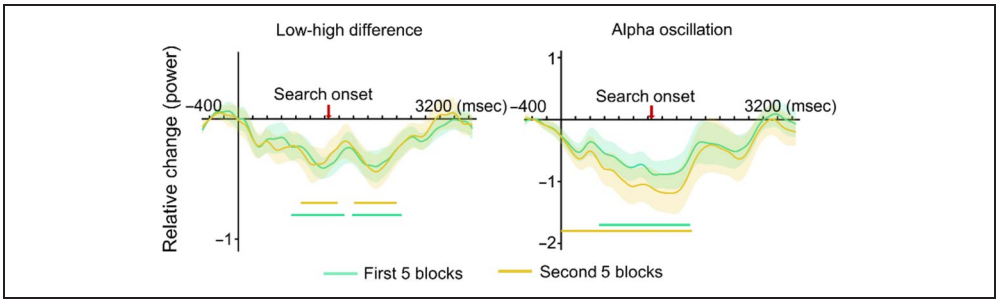


Figure 3. The low–high difference (left) and alpha oscillation (right) for the first and second five blocks in the learning condition.

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn_a_61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

Because we consistently administered the baseline condition before the learning condition, to avoid the potential influences from learned suppression, one might question whether the observed results above were attributable to factors such as fatigue, practice effects, or any consequences associated with the fixed testing order. To address this concern, we partitioned the 10 blocks in the learning condition into the first five and second five blocks. If the observed effects were linked to the testing order, differences between the first five and second five blocks would be anticipated. However, no such distinctions were observed for alpha oscillation or the low–high difference (see Figure 3). Notably, we also applied an inverted encoding model (Foster, Sutterer, Serences, Vogel, & Awh, 2016) to track the processing of target and distractor locations. Consistent with previous studies (e.g., van Moorselaar, Lampers, Cordesius, & Slagter, 2020), our decoding results revealed that, through alpha activity, we could reliably track the target location but not the distractor location (see Figure S2 in the Supplemental Material).

DISCUSSION

The present study aimed to investigate the neural mechanisms underlying the learned suppression of salient distractors using the frequency tagging technique (Malinowski et al., 2007; Ding et al., 2006; Müller et al., 2003; Morgan et al., 1996). Behaviorally, participants demonstrated proficient learned suppression, as evidenced by slower RTs for distractors presented in the high-probability location compared to low-probability locations. Importantly, as observed in tagging responses and alpha oscillation, the results unveiled the anticipatory neural mechanism underpinning learned suppression. This was evidenced by stronger tagging responses toward the high-probability location relative to low-probability locations, as well as a general reduction in alpha power before the search onset. In addition, the ability to differentiate the high-probability location from other low-probability locations before the search onset was primarily observed in participants with better learning performance.

It is generally agreed that learned suppression acts on the spatial priority map, which ultimately determines attentional selection in an all-or-nothing manner (Duncan, van Moorselaar, & Theeuwes, 2023; Ferrante et al., 2018; Sprague & Serences, 2013; Fecteau & Munoz, 2006). This map represents a topographic space encoding the priority of individual locations by integrating signals from sensory input, current goals, and SL (Theeuwes et al., 2022; Zelinsky & Bisley, 2015). According to this framework, if locations are likely to contain a target, the location is upregulated within the spatial priority map, whereas locations with a higher probability of containing distracting information are downregulated, resulting in the observed learned suppression.

Note that, in the present study, we did not employ tagging for the search elements; instead, the placeholders (i.e., the stimulus background) were tagged continuously throughout the trial, with participants explicitly instructed to ignore them. Unlike traditional frequency tagging research, which typically involves continuously tracking attention at a specific location (Ding et al., 2006; Morgan et al., 1996), our experiment required participants to constantly shift their attention among four locations. Consequently, this setup led to a disconnect between the stimuli and their background. In other words, the background (i.e., tagging placeholders) served as a distraction for the search task. Previous studies have shown that enhancing attention to a stimulus causes the suppression of unattended (distracting) stimulus simultaneously (Andersen & Müller, 2010; Müller & Hübner, 2002), thereby leading to reduction in frequency power when tagging the unattended stimulus. This may explain the negative power observed for tagging frequencies in our study (Figure 2A), as the tagging placeholders were disregarded throughout the trial.

Moreover, lower negative tagging responses to placeholders indicate that more attentional resources were allocated to suppressing placeholders (Andersen & Müller, 2010; Müller & Hübner, 2002). Because, at any moment in time, the total amount of attentional resources is limited, allocating fewer resources to processing search elements leaves more resources available for suppressing placeholders. Thus, lower negative tagging responses to placeholders reflect fewer attentional resources directed toward processing search elements, a reversed pattern compared to processing placeholders. In the learning condition, compared to the baseline condition, participants likely allocated less attention to the task, as they had already learned the high-probability distractor location, resulting in more attention to processing placeholders by showing lower tagging responses. This was further evident in behavioral findings, where overall mean RTs in the learning condition (1057 msec) were significantly faster than those in the baseline condition (1116 msec), $p < .001$, suggesting that the learning condition was generally easier for participants, requiring less attentional focus. Accordingly, the lower tagging frequency power observed at low-probability locations relative to high-probability locations (Figure 2C) indicates that participants allocated fewer attentional resources to processing stimuli at low-probability locations. Conversely, participants directed more attentional focus to stimuli presented at the high-probability location before the search onset. This anticipatory focus likely facilitated the suppression of the high-probability location immediately after search initiation.

This finding appears to contradict earlier studies suggesting that learned suppression operates proactively, without requiring prior attentional deployment to distractor locations (Huang, Donk, & Theeuwes, 2022; Kong, Li, Wang, & Theeuwes, 2020; Gaspelin, Leonard, & Luck, 2015, 2017). For instance, Huang et al. (2022) demonstrated that, in the context of SL, RTs to probes at high-probability

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn_a_61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

distractor locations were already suppressed before display onset, consistent with proactive suppression. In contrast, more recent evidence challenges this view. Using a variant of the search-probe paradigm similar to that of Chang, Dube, Golomb, and Leber (2023), Huang et al. (2022) demonstrated that the high-probability location was suppressed relative to all other locations. However, despite this suppression during search, they observed enhanced—rather than reduced—probe discrimination accuracy at this location in the probe task. This suggests that the high-probability location was initially attended before it was suppressed, supporting a reactive suppression mechanism (e.g., Chen, Zhang, Li, Jensen, & Theeuwes, 2025; Chang et al., 2023; Makovski, 2019; Won, Kosoyan, & Geng, 2019; Beck, Luck, & Hollingworth, 2018; Moher & Egeth, 2012).

Furthermore, Chen et al. (2025) employed the same task as we have used here while concurrently measuring microsaccades, a well-established marker of covert attention. Consistent with the current and previous findings (Wang & Theeuwes, 2018a, 2018b, 2018c), they found that the high-probability distractor location was suppressed relative to all other locations. Yet, this learned suppression was accompanied by increased microsaccade rates before stimulus onset, directed toward the high-probability location. This indicates that covert attention was initially allocated to the distractor location before suppression occurred, providing strong evidence for a reactive suppression mechanism. Taken together, these findings consistent with the present results suggest that covert attention is first allocated to distractor locations before suppression is engaged, reinforcing the view that suppression mechanisms may operate reactively rather than purely proactively.

Using EEG recordings, Wang et al. (2019) observed a relative increase in presearch alpha power for electrodes contralateral to the high-probability location, providing direct neural evidence for proactive learned suppression as well. However, this observation contrasts with findings from other related studies (Ferrante, Zhigalov, Hickey, & Jensen, 2023; van Moorselaar et al., 2020; van Moorselaar & Slagter, 2019; Noonan et al., 2016). For instance, a recent magnetoencephalography study showed that prestimulus neural excitability was reduced in the lateralized early visual cortex associated with the high-probability distractor locations, indicating the recruitment of proactive attentional mechanisms (Ferrante et al., 2023). However, Ferrante et al. did not observe alpha activity linked to learned suppression. A possible explanation for this discrepancy is that the learned suppression mechanism may require more complex processing to effectively suppress upcoming distractors in more demanding search tasks, which could involve alpha oscillation to support learned suppression. This may not apply when search is relatively easy. Indeed, previous research indicates that the modulation of alpha power, indicative of distractor suppression, is influenced by the perceptual load (Gutteling, Sillekens, Lavie, & Jensen, 2022). Ferrante et al. (2023) adopted a rapid, invisible frequency tagging method in a search task that was comparatively less challenging, potentially accounting for the lack of alpha oscillation effects observed in their study. Although the current study observed proactive alpha activity, its role may differ from that of tagging responses associated with learned suppression. This distinction highlights the need for further research to clarify the specific contributions of alpha oscillations and tagging responses in learned suppression mechanisms.

In summary, our findings provide direct neural evidence supporting the necessity of attending to distractor locations before suppression, indicating a reactive suppression mechanism.

Corresponding authors: Xinyu Li, Department of Psychology, Zhejiang Normal University, Yingbin Avenue 688, Jinhua, China 321000, e-mail: xyli@zjnu.cn or Benchi Wang, Institute for Brain Research and Rehabilitation, South China Normal University, Zhongshan Road West 55, Guangzhou, China 510000, e-mail: wangbenchi.swift@gmail.com.

Data Availability Statement

The code and data we used can be accessed through https://github.com/zhouyinfei2025/Shared_code. Supplemental Material can be accessed on this article's homepage: http://doi.org/10.1162/JOCN.a.61.

Author Contributions

Y. Z. and L. W. equally contributed to the current research. L. W., X. L., and B. W. designed the experiment. Y. Z., L. W., and B. W. collected and analyzed the data. All authors wrote the article and approved the final version of the manuscript for submission and declared no competing interests.

Funding Information

This research was supported by the Natural Science Foundation of Guangdong (https://dx.doi.org/10.13039/501100003453), grant (2023A1515012789) to B. W.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn.a.61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

REFERENCES

Andersen, S. K., & Müller, M. M. (2010). Behavioral performance follows the time course of neural facilitation and suppression during cued shifts of feature-selective attention. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 13878–13882. https://doi.org/10.1073/pnas.1002436107, PubMed: 20643918

Antonov, P. A., Chakravarthi, R., & Andersen, S. K. (2020). Too little, too late, and in the wrong place: Alpha band activity does not reflect an active mechanism of selective attention. *Neuroimage*, *219*, 117006. https://doi.org/10.1016/j.neuroimage.2020.117006, PubMed: 32485307

Beck, V. M., Luck, S. J., & Hollingworth, A. (2018). Whatever you do, don't look at the…: Evaluating guidance by an exclusionary attentional template. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 645–662. https://doi.org/10.1037/xhp0000485, PubMed: 29035075

Chang, S., Dube, B., Golomb, J. D., & Leber, A. B. (2023). Learned spatial suppression is not always proactive. *Journal of Experimental Psychology: Human Perception and Performance*, *49*, 1031–1041. https://doi.org/10.1037/xhp0001133, PubMed: 37199949

Chen, S., Zhang, X., Li, X., Jensen, O., Theeuwes, J., & Wang, B. (2025). Statistical learning drives anticipatory micro-saccades toward suppressed distractor locations. *bioRxiv*. https://doi.org/10.1101/2025.05.11.652318

Ding, J., Sperling, G., & Srinivasan, R. (2006). Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency. *Cerebral Cortex*, *16*, 1016–1029. https://doi.org/10.1093/cercor/bhj044, PubMed: 16221931

Duncan, D. H., van Moorselaar, D., & Theeuwes, J. (2023). Pinging the brain to reveal the hidden attentional priority map using encephalography. *Nature Communications*, *14*, 4749. https://doi.org/10.1038/s41467-023-40405-8, PubMed: 37550310

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*, 382–390. https://doi.org/10.1016/j.tics.2006.06.011, PubMed: 16843702

Ferrante, O., Patacca, A., Di Caro, V., Della Libera, C., Santandrea, E., & Chelazzi, L. (2018). Altering spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex*, *102*, 67–95. https://doi.org/10.1016/j.cortex.2017.09.027, PubMed: 29096874

Ferrante, O., Zhigalov, A., Hickey, C., & Jensen, O. (2023). Statistical learning of distractor suppression downregulates prestimulus neural excitability in early visual cortex. *Journal of Neuroscience*, *43*, 2190–2198. https://doi.org/10.1523/jneurosci.1703-22.2022, PubMed: 36801825

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology*, *115*, 168–177. https://doi.org/10.1152/jn.00860.2015, PubMed: 26467522

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*, 154. https://doi.org/10.3389/fpsyg.2011.00154, PubMed: 21779269

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychological Science*, *26*, 1740–1750. https://doi.org/10.1177/0956797615597913, PubMed: 26420441

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2017). Suppression of overt attentional capture by salient-but-irrelevant color singletons. *Attention, Perception, & Psychophysics*, *79*, 45–62. https://doi.org/10.3758/s13414-016-1209-1, PubMed: 27804032

Goschy, H., Bakos, S., Müller, H. J., & Zehetleitner, M. (2014). Probability cueing of distractor locations: Both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in Psychology*, *5*, 1195. https://doi.org/10.3389/fpsyg.2014.01195, PubMed: 25414676

Gutteling, T. P., Sillekens, L., Lavie, N., & Jensen, O. (2022). Alpha oscillations reflect suppression of distractors with increased perceptual load. *Progress in Neurobiology*, *214*, 35533812. https://doi.org/10.1016/j.pneurobio.2022.102285, PubMed: 35533812

Huang, C., Donk, M., , & Theeuwes, J. (2022). Proactive enhancement and suppression elicited by statistical regularities in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *48*, 443–457. https://doi.org/10.1037/xhp0001002, PubMed: 35324244

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186. https://doi.org/10.3389/fnhum.2010.00186, PubMed: 21119777

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, *37*, 43–75. https://doi.org/10.1146/annurev.ps.37.020186.000355

Kerzel, D., Balbiani, C., Rosa, S., & Huynh Cong, S. (2022). Statistical learning in visual search reflects distractor rarity, not only attentional suppression. *Psychonomic Bulletin & Review*, *29*, 1890–1897. https://doi.org/10.3758/s13423-022-02097-x, PubMed: 35445289

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*, 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003, PubMed: 16887192

Kong, S., Li, X., Wang, B., & Theeuwes, J. (2020). Proactively location-based suppression elicited by statistical learning. *PLoS One*, *15*, e0233544. https://doi.org/10.1371/journal.pone.0233544, PubMed: 32479531

Leber, A. B., Gwinn, R. E., Hong, Y., & O'Toole, R. J. (2016). Implicitly learned suppression of irrelevant spatial locations. *Psychonomic Bulletin & Review*, *23*, 1873–1881. https://doi.org/10.3758/s13423-016-1065-y, PubMed: 27225635

Lin, R., Meng, X., Chen, F., Li, X., Jensen, O., Theeuwes, J., et al. (2024). Neural evidence for attentional capture by salient distractors. *Nature Human Behaviour*, *8*, 932–944. https://doi.org/10.1038/s41562-024-01852-5, PubMed: 38538771

Makovski, T. (2019). Preparing for distraction: Attention is enhanced prior to the presentation of distractors. *Journal of Experimental Psychology: General*, *148*, 221–236. https://doi.org/10.1037/xge0000509, PubMed: 30346200

Malinowski, P., Fuchs, S., & Müller, M. M. (2007). Sustained division of spatial attention to multiple locations within one hemifield. *Neuroscience Letters*, *414*, 65–70. https://doi.org/10.1016/j.neulet.2006.12.001, PubMed: 17207932

Moher, J., & Egeth, H. E. (2012). The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Attention, Perception, & Psychophysics*, *74*, 1590–1605. https://doi.org/10.3758/s13414-012-0358-0, PubMed: 22893004

Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences, U.S.A.*, *93*, 4770–4774. https://doi.org/10.1073/pnas.93.10.4770, PubMed: 8643478

Müller, M. M., & Hübner, R. (2002). Can the spotlight of attention be shaped like a doughnut? Evidence from steady-state visual evoked potentials. *Psychological Science*, *13*, 119–124. https://doi.org/10.1111/1467-9280.00422, PubMed: 11933994

Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight.

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn_a_61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025

Nature, *424*, 309–312. https://doi.org/10.1038/nature01812, PubMed: 12867981

Noonan, M. P., Adamian, N., Pike, A., Printzlau, F., Crittenden, B. M., & Stokes, M. G. (2016). Distinct mechanisms for distractor suppression and target facilitation. *Journal of Neuroscience*, *36*, 1797–1807. https://doi.org/10.1523/JNEUROSCI.2133-15.2016, PubMed: 26865606

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, *15*, 4. https://doi.org/10.1167/15.6.4, PubMed: 26024451

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., et al. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *European Journal of Neuroscience*, *22*, 2917–2926. https://doi.org/10.1111/j.1460-9568.2005.04482.x, PubMed: 16324126

Sauter, M., Liesefeld, H. R., Zehetleitner, M., & Müller, H. J. (2018). Region-based shielding of visual search from salient distractors: Target detection is impaired with same- but not different-dimension distractors. *Attention, Perception, & Psychophysics*, *80*, 622–642. https://doi.org/10.3758/s13414-017-1477-4, PubMed: 29299850

Sprague, T. C., & Serences, J. T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*, *16*, 1879–1887. https://doi.org/10.1038/nn.3574, PubMed: 24212672

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606. https://doi.org/10.3758/bf03211656, PubMed: 1620571

Theeuwes, J., Bogaerts, L., & van Moorselaar, D. (2022). What to expect where and when: How statistical learning drives visual selection. *Trends in Cognitive Sciences*, *26*, 860–872. https://doi.org/10.1016/j.tics.2022.06.001, PubMed: 35840476

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552–564. https://doi.org/10.1037/0096-3445.134.4.552, PubMed: 16316291

van Moorselaar, D., Lampers, E., Cordesius, E., & Slagter, H. A. (2020). Neural mechanisms underlying expectation-dependent inhibition of distracting information. *eLife*, *9*, e61048. https://doi.org/10.7554/eLife.61048, PubMed: 33320084

van Moorselaar, D., & Slagter, H. A. (2019). Learning what is irrelevant or relevant: Expectations facilitate distractor inhibition and target facilitation through distinct neural mechanisms. *Journal of Neuroscience*, *39*, 6953–6967. https://doi.org/10.1523/JNEUROSCI.0593-19.2019, PubMed: 31270162

Wang, B., & Theeuwes, J. (2018a). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, & Psychophysics*, *80*, 860–870. https://doi.org/10.3758/s13414-018-1493-z, PubMed: 29476331

Wang, B., & Theeuwes, J. (2018b). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 13–17. https://doi.org/10.1037/xhp0000472, PubMed: 29309194

Wang, B., & Theeuwes, J. (2018c). Statistical regularities modulate attentional capture independent of search strategy. *Attention, Perception, & Psychophysics*, *80*, 1763–1774. https://doi.org/10.3758/s13414-018-1562-3, PubMed: 29968080

Wang, B., & Theeuwes, J. (2020). Salience determines attentional orienting in visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, *46*, 1051–1057. https://doi.org/10.1037/xhp0000796, PubMed: 32757594

Wang, B., van Driel, J., Ort, E., & Theeuwes, J. (2019). Anticipatory distractor suppression elicited by statistical regularities in visual search. *Journal of Cognitive Neuroscience*, *31*, 1535–1548. https://doi.org/10.1162/jocn_a_01433, PubMed: 31180265

Won, B.-Y., Kosoyan, M., & Geng, J. J. (2019). Evidence for second-order singleton suppression based on probabilistic expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 125–138. https://doi.org/10.1037/xhp0000594, PubMed: 30596437

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences*, *1339*, 154–164. https://doi.org/10.1111/nyas.12606, PubMed: 25581477

Zhigalov, A., Herring, J. D., Herpers, J., Bergmann, T. O., & Jensen, O. (2019). Probing cortical excitability using rapid frequency tagging. *Neuroimage*, *195*, 59–66. https://doi.org/10.1016/j.neuroimage.2019.03.056, PubMed: 30930309

Downloaded from http://direct.mit.edu/jocn/article-pdf/37/12/2493/2530392/jocn.a.61.pdf by UNIVERSITEIT LEIDEN user on 15 December 2025