Tracking neural markers of template formation and implementation in attentional inhibition under different distractor consistency

Wen Wen[1], Zhibang Huang[1], Yin Hou[1], Sheng Li[*1,2,3,4]

1. School of Psychological and Cognitive Sciences, Peking University, Beijing, China
2. Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China
3. PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China
4. Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

This work was supported by a grant from National Key R&D Program of China (2017YFB1002503).

Correspondence concerning this article should be addressed to Sheng Li, School of Psychological and Cognitive Sciences, Peking University, 5 Yiheyuan Road, Haidian, Beijing, 100871, China. E-mail: sli@pku.edu.cn

Word count: 5509

Abstract

To perform challenging visual search tasks, attention should be selectively allocated to promote targets and inhibit distractors. Rejection templates based on the distractor's feature can be built to constrain the search process by filtering-out matched stimuli. We measured electroencephalography (EEG) of human participants when they performed a visual search task under conditions in which the distractor cues were constant within a block (fixed-cueing) or changed on a trial-by-trial basis (varied-cueing). Sustained representation of the cued colors could be decoded from the whole-brain EEG signal during the retention interval in the fixed-cueing condition, implying a formed rejection template. By contrast, transient representation of the cued colors was observed in the varied-cueing condition only after cue onset with significant but lower decoding accuracy than the fixed-cueing condition. The differential neural representations of the to-be-ignored color in the two cueing conditions agreed with the visual search performance that the fixed-cueing condition demonstrated a larger inhibition benefit of the distractor cueing than the varied-cueing condition. During the retention interval, we also observed stronger posterior alpha lateralization and mid-frontal theta/beta power in the varied-cueing condition, indicating the cognitive costs in template formation caused by the trialwisely change of distractor colors. The stable template formed in the fixed-cueing condition was utilized to inhibit the matched distractors in the search array and this was accompanied by larger mid-frontal theta activity. Taken together, our findings highlighted the critical roles of distractor consistency in linking template formation to successful distractor inhibition and the functional relevance of frontal theta activity in support of template formation and implementation.

Keywords: rejection template, attentional suppression, visual search, cognitive control, theta oscillation.

Introduction

Human beings can allocate limited attentional resources to task-relevant items whilst filtering out irrelevant distractors. Extensive evidence has demonstrated that target templates held in memory facilitate visual performance [1–5]. Importantly, the neural representation of target templates during the preparation stage is positively correlated with subsequent task performance [5–7]. By contrast, divergent findings have been reported regarding whether foreknowledge of the to-be-ignored feature, referred to as a rejection template [8,9], can benefit search efficiency. For those studies that showed the behavioral suppression benefits of the rejection template, the emergence of such benefits seemed to be dependent on multiple factors [10–16]. Of note, the consistency of the to-be-ignored feature across trials was particularly prepotent [17–20]. These studies have demonstrated robust benefits of the distractor cues when the cues remained unchanged within a block (fixed-cueing condition) and minimum benefits if not absent when the cues changed on a trial-by-trial basis (varied-cueing condition).

To reconcile the discrepancies of findings in the fixed and varied distractor-cueing conditions, researchers proposed that a stable rejection template is essential in distractor inhibition and can be created only under the fixed-cueing condition, whereas the varied-cueing context obstructed its formation [8,17,21]. So far, there was limited neural evidence shedding light on the existence of rejection templates. Reeder and colleagues [22] provided their participants with three types of cues (positive cues, negative cues, and neutral cues) and found that negative cues that indicated the distractor feature led to decreased responses in the early visual cortex (V1-V3) prior to the search array. However, the authors later reported that this decrement in the early visual cortex was not a feature-specific suppression and there

were no distinctive representations of the rejection templates [23]. Therefore, it is necessary to examine whether selective representations of the rejection templates could be identified.

Recent studies come to reach the consensus that distractor inhibition could be learned when statistical regularities of distractors' feature [24] or spatial location [25–27] are expected (see Geng et al., 2019 for a review). By adopting a variant of Posner cueing paradigm, Noonan et al. [20] found that benefits of distractor suppression, as indexed by a decreased amplitude of early visual component P1, emerged only in the blocked condition where the distractor's location was fixed. Similarly, by manipulating distractors' locations, van Moorelaar and Slageter [29] showed that distractor-location-learning induced behavioral benefits appeared after several repetitions and the learned expectation of distractors reduced stimulus-specific processing as reflected by the decreased Pd component and location-decoding accuracy. On the other hand, Vatterott and Vecera [19] examined the experience-dependent nature of feature-based distractor suppression by manipulating the color contingency of a singleton distractor in each trial within and between blocks. The capture effect of the singleton distractor decreased within a block where the singleton's color remained unchanged but rebounded when a new block started with another colored singleton. With all the evidence showing the influence of distractor consistency on stimulus processing, yet it remained less understood how distractor consistency modulates the process of template formation and subsequent implementation of the template for distractor filtering.

Mid-frontal theta-band activities have been linked to goal-related prioritization in working memory tasks [30,31] and conflict solving in the framework of cognitive control [32,33]. Specifically, stronger theta activity was evoked by distractor cues relative to target cues during the delay period as a means of reducing interference from matched distractors [34]. Meanwhile, Dugué and colleagues [35] observed stronger post-stimulus theta activity in correct

trials than incorrect trials in a visual search task, revealing the association between post-stimulus theta power and target selection. Although previous studies have indicated the tight connection between template-based distractor inhibition and cognitive control [12,36], the exact functional roles of theta activity in distractor inhibition, especially pertaining to template formation and implementation, need more empirical investigations.

The current study aimed to reveal the feature-specific representation of rejection template and the associated neural dynamics. Participants performed a cued visual search task in which they were asked to locate the target and discriminate the orientation of the bar inside of the target (Figure 1A). The participants were provided with distractor cues and were informed that the search items in the same color as the distractor cues would never be the target. The experiment contained the fixed-cueing and varied-cueing conditions that were defined according to the consistency of the distractor cues within a block of trials. We recorded electroencephalography (EEG) signals when the participants performed the task and applied multivariate decoding method on the EEG signals to reveal the neural representations of the to-be-ignored colors during the retention interval. We also examined the oscillatory activities at different frequency bands to identify the neural markers that corresponded to attentional selection, memory transformation, and cognitive control in the template formation and implementation stages. Our results provide novel neural evidence in support of the crucial roles of distractor consistency in linking template formation to successful distractor inhibition.

Method

Participants.

We tested thirty-four participants in total (mean age = 21.3, SD = 2.2), ranging from 18 to 27 years old, 15 females, 1 left-handed). According to the G-power estimation [37], thirty

participants at least should be tested to guarantee a medium effect size (f = 0.25) and a high statistic power (90%). Participants were compensated with monetary rewards. All participants had normal or corrected-to-normal (color) vision and had no history of psychiatric or neurological disorders. All participants provided written informed consent prior to the experiment. The study was approved by the Committee for Protecting Human and Animal Subjects at the School of Psychological and Cognitive Sciences at Peking University (IRB protocol #2017-11-02).

Stimuli and procedure.

Participants were faced with a CRT monitor (refresh rate = 60 Hz) in a dark shielded EEG booth and their heads were stabilized using a chin rest. The schematic procedure of a trial is shown in Figure 1A. After a jittered blank (~1.5 s), a cue patch (1.5° × 1.5°) and an equal-sized gray patch were presented (1.5° to the center) for 0.1 s. The gray patch was presented to equalize the visual input in both hemispheres. The cue patch could appear at either hemifield with equal probability. Participants were instructed that target in the following search array would never be in the same color as the cued color. The color cue remained the same within one block in the fixed-cueing condition while it changed trialwisely in the varied-cueing condition. Two gray frames of the same size and locations as the two patches were showed during the retention interval (1.5 s). The search array contained eight items of different shapes (diamond/circle/hexagon, in the same square size of 2°× 2°) and colors (four in target color and four in distractor color). Participants were instructed to find the target, defined solely by its shape (diamond), and discriminate the orientation of the bar inside of it. The target's color was randomly selected in each trial. In matched trials (75%), the distractor's color was the same as the cued color. In neutral trials (25%), the distractor's color was a new color different from the target color and the cued color. The

search array lasted for 0.2 s. A final memory test required participants to select the patch that was in the same color as the cue from two presented patches. Both the search task and the memory test had a response limit of 2 s.

Each cueing condition (fixed and varied) had 12 blocks with 64 trials per block. Six cue colors (selected from L*a*b color space, L = 70, equally distributed in the a*b plane) appeared in the first six blocks as well as the last six blocks in the fixed-cueing condition. Participants came twice and completed the fixed-cueing and varied-cueing conditions on separate days. The order of the conditions was counterbalanced between participants. Participants were instructed to look at the fixation dot and their eye movements were monitored using Eyelink 1000/1000Plus (1000 Hz, monocular).

EEG recording.

EEG data were acquired at a sampling rate of 1000 Hz using BrainAmps amplifier from a 64-channel EasyCap (Brain Products, Munich, Germany) with Ag/AgCl electrodes placed according to the 10-20 system. An external electrode placed below the right eye was used to measure eye movements. Electrode impedance was kept below 10 kΩ.

Data analysis.

Behavioral data: Only trials with correct responses in both the search task and the memory task were included for search task's reaction times (RTs) analysis. Besides, we recursively removed outlier trials where search RTs were beyond 2.5 standard deviations from the mean. This resulted in 4.4% of the trials in the fixed-cueing condition and 3.7% of the trials in the varied-cueing condition being discarded. There was a general practice effect given that participants completed the two cueing conditions on separate days. Hence, we calculated the benefit index for each cueing condition to represent the RT facilitation resulted from the inhibition of matched distractors:

$$\text{RT benefit index} = 100 * \frac{RT_{neutral} - RT_{matched}}{(RT_{neutral} + RT_{matched})/2}$$

To examine the possible influence of testing order on the suppression benefit, we performed a two-way mixed ANOVA on benefit index with testing order (fixed-cueing first or varied-cueing first) as between-subject factor and cueing condition (fixed vs. varied) as within-subject factor. We further split the trials into five bins based on the search RT percentile for each trial type in each cueing condition. Then, we calculated the benefit index for each percentile bin in each cueing condition and conducted one-sample t-tests (two-tailed, alpha = .05) to compare the obtained benefit indices with zero. A two-way repeated ANOVA with factors of cueing condition (fixed vs. varied) and trial type (matched vs. neutral) was performed on the accuracies of the search task and the memory test. We reported the corrected degrees of freedom (Greenhouse-Geisser correction) when the equal variances assumption was violated.

EEG preprocessing: Offline preprocessing was performed with EEGLAB [38]. After down-sampled to 500 Hz, the data was bandpass-filtered (1~30 Hz). Bad channels were spherically interpolated. We then re-referenced the data to the averaged signal of mastoids and performed independent component analysis (ICA) to correct ocular movements and other artifacts. Finally, the continuous EEG was epoched for further analyses (the exact epoch length and time-locked onset point were chosen depending on specific analyses). Besides behavioral exclusion described above, epochs in which any channel showed above-threshold amplitude (60 μV) at any time point of the epoch were removed. Since matched trials and neutral trials differed from each other only after the onset of search array, we used both types of trials when the analysis focused on the activities during the retention interval.

EEG decoding: To reveal the neural representations of the rejection templates, we performed color-decoding based on the whole-brain electrodes' broad-band EEG data (EOG

excluded) during the retention interval. The decoding was separately performed on cue-left and cue-right trials and the averaged results were used for statistical analysis. The continuous EEG was segmented into -0.2~1.6 s relative to the cue onset and baseline-corrected to the mean amplitude of the pre-stimulus period. We Gaussian-smoothed (window size = 16 ms) the data along the time dimension to reduce temporal noises. At each time point, data was partitioned into six training folds and one test fold. To create an unbiased classifier, the number of trials for each cue color in the training set was equalized by subsampling. Trial-averaged training set was generated by averaging the trials from the same cue color. We first estimated the covariance matrix based on the trial-averaged training set. Next, we computed the pair-wise Mahalanobis distances between test trials and the trial-averaged training set using the covariance matrix [39]. A smaller distance toward the trial-averaged activation of one specific color cue condition indicated a larger pattern similarity. Therefore, the decoding was marked as a success when the distance between this given trial and the activation profile of the same color was smaller than any of other colors' activation profile in the training data. Cross-validation was carried out in a leave-one-fold-out manner. This routine was iterated 1000 times and trials were randomly assigned to the training and test set in each iteration. The averaged decoding accuracy was taken for the statistical analysis. The accuracy time series were Gaussian-smoothed (window size = 16 ms) and compared with the chance level (i.e., 1/6). Cluster-based permutation was performed on the decoding accuracy time-series over participants (alpha = .05, cluster-based nonparametric alpha = .05, cluster statistic = sum, two-tail, permutation times = 10000 [40]).

Cross-temporal decoding: We followed the same decoding procedure as explained above except that we trained the model using data at one time point and tested it on all time points. We down-sampled the data into 50 Hz for computational efficiency. For the 2D result map

(as in Figure 2B), the diagonal line represents the accuracies when training and testing were performed at the same time point while the off-diagonal points correspond to the temporal-generalized decoding. We adopted the leave-one-fold-out cross-validation method throughout the decoding process. Hence, the training-fold and the test-fold were never based on the same trials no matter whether the training and test time points were the same or different. This procedure was iterated 1000 times for each participant and the averaged result was used for group statistical tests. Cluster-based permutation test was conducted on the 2D result map at the participants' population level (alpha = .01, cluster-based nonparametric alpha = .01, cluster statistic = sum, two-tail, permutation times = 10000 [40]).

Single-trial's averaged accuracy: Single-trial's averaged accuracy of the matched trials was calculated to assess the relationship between the representation of the to-be-ignored color and the following search performance. Instead of calculating the averaged decoding accuracy (i.e., percentage of correct prediction) for each time point of the test-fold, we applied the trained model to each trial in the test-fold and labeled the prediction of the trial using a binary tag (1 for correct, 0 for incorrect). Once this was done for all time points of the trial, we took the averaged binary tags (0~1600 ms) as the averaged accuracy for this specific trial. This procedure was iterated for 1000 times and the mean accuracy was used for statistics. We calculated the Pearson's correlation between trials' averaged accuracy and search RTs for each participant. The populations' correlation coefficients in each cueing condition were tested against zero using the one-sample bootstrapping t-test (alpha = 0.05, two-tail, bootstrap samples = 10000).

Posterior alpha lateralization: A surface Laplacian filter (a 10th-order Legendre polynomial and a lambda of 10^{-5}) was applied to increase topographical specificity and reduce the effects of volume conduction (Cohen, 2019). To calculate the posterior lateral

alpha power, we segmented the continuous data into -0.4~1.6 s relative to the cue onset, and the averaged ERP was subtracted from each trial to remove phase-locked activity evoked by stimulus onset. Hilbert transform was used to extract the alpha-band (8~12 Hz) power of each trial using posterior occipital electrodes (P5/6, P7/8, PO7/8). Power data was decibel normalized to the averaged baseline (-0.4~-0.1 s). This process was separately performed on left-side cue and right-side cue trials and the averaged results were used for the statistical analysis. We calculated the alpha lateralization as the power difference between contralateral and ipsilateral electrodes (contralateral minus ipsilateral).

Time-frequency analysis: We performed time-frequency decomposition on the Laplacian-filtered data using Fieldtrip [42]. The continuous data were epoched -1.5~3 s relative to cue onset. After the time-frequency decomposition, we selected data between -0.4~1.6 s for further analysis. Each trial's signal was decomposed using Morlet wavelet convolution every 50 ms. The frequencies ranged from 1~30 Hz in 30 linearly spaced steps and the width of the wavelet kernel increased linearly from 1 to 10 along with frequency bins. Normalization was done by decibel baselined to the pre-stimulus interval (-0.4~-0.1 s). Time-frequency decomposition was separately done on the left- and right-cued trials. We further combined the resulting data from the left- and right-cued trials so that the left-hemifield electrodes were ipsilateral to the stimulus. To test the difference between conditions, cluster-based permutation was performed at the group level. Adjacent time-frequency time points exceeding the threshold (alpha = .01, two-tail) were grouped as a cluster. The cluster-level statistic was calculated by taking the sum of the difference values within the cluster (cluster alpha = .05). The number of random permutations using the Monte Carlo method was set to 10000.

We also examined the neural oscillations related to template implementation during visual search. Continuous data were epoched -1~2 s relative to search array onset. We adopted the same time-frequency decomposition method to obtain the oscillatory activities during -0.4~1.2 s for all trials.

Results

Behavioral suppression benefits in matched trials

We observed larger response benefits in the fixed-cueing condition than the varied-cueing condition (F(1,32) = 9.661, p = .004, η_p^2 = .232, Figure 1B) but no significant effect of testing order (F(1,32) = 1.558, p = .221, η_p^2 = .046). There was no significant interaction effect between the two factors (F(1,32) = .064, p = .802, η_p^2 = .002)[1].

As shown in Figure 1C, we analyzed the benefit index across five RT percentile bins. The results showed that the suppression benefits were evident across all bins in the fixed-cueing condition (all p < .001). However, in the varied-cueing condition, we observed significant suppression benefit only in the percentile bin with longest RTs (p < .001). The early and robust behavioral benefits in the fixed-cueing condition indicated the engagement of proactive inhibition when the to-be-ignored color remained the same. By contrast, the benefit of reactive inhibition was relatively late and emerged only in the group of trials with longest RTs in the varied-cueing condition. These results were consistent with our previous findings that cognitive control played a crucial role in distractor inhibition in the varied-cueing condition [18].

We further analyzed the results of accuracy for visual search and memory test (Figure

[1] See Supplementary Materials Figure S1 for the search RTs of different conditions.

1D). For the visual search accuracy, there were no significant differences between conditions (F(1,33) = 1.784, p = .191, η_p^2 = .051) or trial types (F(1,33) = 3.115, p = .087, η_p^2 = .086) but their interaction effect was significant (F(1,33) = 8.983, p = .005, η_p^2 = .214). Simple effect analysis showed that participants were more accurate in the matched trials than the neutral trials in the fixed-cueing condition (F(1,33) = 11.31, p = .002). However, this difference was not evident in the varied-cueing condition (F(1,33) = 0.19, p = .662). For the memory test, participants were more accurate in the fixed-cueing condition (F(1, 33) = 100.318, p < .001, η_p^2 = .752) and the matched trials (F(1, 33) = 93.234, p < .001, η_p^2 = .739). The significant interaction effect (F(1, 33) = 46.202, p < .001, η_p^2 = .583) indicated that when the cued color reappeared in the search array as the distractor color, it promoted memory performance to a larger extent in the varied-cueing condition (5.8%, F(1, 33) = 83.26, p < .001) than in the fixed-cueing condition (1.3%, F(1, 33) = 19.28, p < .001). This result might be explained by the refreshment strategy (or perceptual sampling) that using the reappeared matched distractors to enhance the memory of the to-be-ignored distractor color in the varied-cueing condition [9].

EEG decoding of the to-be-ignored color during retention interval

We performed decoding analysis on EEG data to reveal the representation of the distractor color during the retention interval. As shown in Figure 2A, color information could be decoded in both conditions (fixed-cueing: 0.078~1.6 s, cluster p < .001; varied-cueing: 0.26~0.806 s, cluster p < .001). Importantly, the decoding accuracy was higher in the fixed-cueing condition compared with the varied-cueing condition, especially after the cue onset (cluster p < .001, 0.08~0.438 s) and prior to search array onset (cluster p = .021, 1.096~1.272 s; cluster p = .005, 1.342~1.6 s).

Next, we focused on representation stability and performed cross-temporal decoding separately for the two cueing conditions. Figure 2B shows the two-dimensional maps of decoding accuracies. There was a wide range of generalization for color decoding between time points in the fixed-cueing condition, indicating a more stable representation of the cued color during the retention interval. In contrast, significant decoding performance was observed only when training and testing were performed on the data from the same time points in the varied-cueing condition, suggesting a highly dynamic representation of the cued color. These results were in line with the larger behavioral suppression benefit in the fixed-cueing condition as such benefit would highly rely on the stable representation of the distractor's feature to form an effective rejection template.

To further unfold the relationship between the neural representation of the cued color and distractor inhibition, we calculated the single-trial's averaged accuracy on the matched trials in the two cueing conditions. We then computed the trialwise correlation between the single-trials' mean accuracies and search RTs for each participant. We hypothesized that stable rejection template would facilitate distractor filtering and therefore lead to faster response in the matched trials. If the neural representation instead reflected initial sensory processing, better representation of the distractor cue would lead to behavioral cost. Results in Figure 2C showed that single-trial's averaged accuracy was positively correlated with search RT in the varied-cueing condition (mean of correlation coefficients r = .014, p = .042) confirming that better representation of the color would cause attraction. This effect was not evident in the fixed-cueing condition (mean of correlation coefficients r = .004, p = .562) which might be explained by the ceiling effect of behavioral suppression benefits. These discrepancies suggested distinct neural representations of the to-be-ignored cue in the two cueing contingencies. In fact, when we performed multivariate representation similarity analysis

(RSA) on the topographical data within and between two cueing conditions (see Supplementary Materials Figure S2), the results showed that RSA scores of the within-condition scheme were higher than the between-condition scheme. Thus, although based on the same sensory information, brain signals during the retention interval manifested different topographical distributions under different cueing conditions.

Posterior alpha lateralization during early stage of retention interval

An increase in posterior alpha-band (8~12 Hz) activity has been regarded as a neural marker of attentional inhibition [43–45]. Correspondingly, the decrease of alpha power at the contralateral side of an attended stimulus indicated the release from inhibition (i.e., increase in attention). As shown in Figure 3A, we observed significant alpha lateralization (contralateral minus ipsilateral) induced by the color cue in both cueing conditions (fixed-cueing condition: 0.282~0.74 s, cluster p = .008; varied-cueing condition: 0.266~1.124 s, cluster p = .001). Stronger alpha lateralization (herein, more negative) was found in the varied-cueing condition relative to the fixed-cueing condition (0.338~1.066 s, cluster p = .003), indicating a larger attentional bias towards the color cue at the initial stage of the retention interval.

Mid-frontal theta and beta activities during the late stage of retention interval

Next, we focused on mid-frontal activities in the two cueing conditions (Figure 3B left). The time-frequency difference map of mid-frontal electrodes showed that varied cues induced stronger oscillatory activities in the range of theta-band and beta-band at the later stage of the retention interval in comparison with the fixed cues (0.35~1.6 s, 2~19 Hz, cluster p = .002). Frontal theta oscillation (3~7 Hz) has been regarded as a reflection of top-down control in goal-directed cognitive processes [32]. In two recent studies, stronger mid-frontal theta was observed during the retention interval if the cue indicated the distractor rather than the target [34] or switching rather than repeating the task sets [46]. The enhanced mid-frontal theta

activity was also suggested to be associated with the prevention of interference from the anticipated distractor [34] or the proactive control of task set switching [46]. In our experiment, the varied-cueing condition shared these critical aspects of task structure with these two studies. Therefore, we consider that both the prevention of interference from the anticipated distractor and the proactive control of task set switching could contribute to the observed mid-frontal theta effects during the retention interval.

On the other hand, frontal beta activity (13~30 Hz) has been proposed to reflect a short-lived transition period between representations (e.g., from latent to active) late in the WM delay in the service of task demands [47]. We speculated that the higher mid-frontal beta activity in the varied-cueing condition (Figure 3B, right-top) reflected a stronger engagement in transferring the cueing information from perceptual representation into rejection template that can be used to filter out the matched distractors in the following search task. However, given the lack of literature on the roles of frontal beta in attentional suppression, future experiments with more specific designs are required to test this interpretation.

Mid-frontal theta activity during visual search period

Finally, time-frequency analysis also highlighted the dominant low-frequency oscillatory activity in mid-frontal electrodes during the search process (1~9 Hz, 0.2-1.2 s, cluster $p < .001$, Figure 4A). Specifically, as shown in Figure 4B, mid-frontal theta power during visual search correlated with individuals' RTs in the matched (fixed-cueing: r = -.398, $p = .020$; varied-cueing: r = -.500, $p = .003$) and neutral trials (fixed-cueing: r = -.445, $p = .009$; varied-cueing: r = -.657, $p < .001$). The observed correlations demonstrated that mid-frontal theta reflected a general process of cognitive control that filters out the distractors of different colors and shapes. Participants who utilized stronger control process were faster in localizing the target.

Of particular interest, we looked into the relationship between post-stimulus theta activity and template implementation. If mid-frontal theta also signaled the implementation of the rejection template on top of the general control process in visual search, we would expect to observe stronger mid-frontal theta for the trials (e.g., matched trials) in which the rejection templates were successfully utilized for search. To test this assumption, we conducted a repeated-measures ANOVA on the mid-frontal theta power with cueing condition (fixed vs. varied) and trial type (matched vs. neutral) as within-subject factors. We did not find significant difference for either cueing condition ($F(1,33) = 1.928$, $p = .174$, $\eta^2 = 0.055$) or trial type ($F(1,33) = 0.006$, $p = .939$, $\eta^2 < .001$). However, there was a marginal significant interaction effect between the two factors ($F(1,33) = 3.675$, $p = .064$, $\eta^2 = 0.100$). Simple effect analysis revealed that the fixed-cueing condition had stronger mid-frontal theta power than the varied-cueing condition in the matched trials ($F(1,33) = 6.94$, $p = .013$), but not in the neutral trials ($F(1,33) = 0.09$, $p = .762$). These results were consistent with the observed larger behavioral suppression benefits in the fixed-cueing condition, indicating that mid-frontal theta served as a neural marker not only for the general control process of distractor filtering but also for the implementation of the established rejection template.

Discussion

Distractor inhibition is challenging and context-dependent. Inconsistent results have been reported concerning whether distractors matched with the to-be-ignored feature could be inhibited. The present study identified the neural signature of the rejection template in the fixed-cueing condition as characterized by the sustained decoding of the to-be-ignored color and the stable cross-temporal decoding performance. By contrast, trial-wise updates of distractor cues in the varied-cueing condition hampered the establishment of stable rejection templates. This was reflected as the smaller behavioral suppression benefit and the transient

decoding performance that did not persist in the late stage of the retention interval. Moreover, the results of single-trial's averaged accuracy implied that the representation of distractor cues in the varied-cueing condition would conversely interfere with the following visual search performance. These findings collectively suggested that creating rejection templates is difficult when the distractor consistency is violated.

The specific processes related to template formation and implementation were further revealed by oscillatory activities in different frequency bands. First, we observed stronger posterior alpha lateralization toward the color cue in the varied-cueing condition than the fixed-cueing condition, reflecting the higher necessity of paying attention to the color cue when it changed on a trial-by-trial basis. Second, the posterior alpha lateralization was followed by mid-frontal theta and beta activities in the late stage of the retention interval. Based on previous literature, we suggest that the theta band activity reflected the control processes such as the prevention of the anticipated distractor interference and the proactive control of cue switching [34,46]. The beta band activity, on the other hand, may index the memory transition process that transferred the color cue from perceptual representation to rejection template [47]. Third, increased mid-frontal theta activity was found after the search array onset and was behaviorally relevant with search performance. This post-stimulus mid-frontal theta activity was stronger in the fixed-cueing condition than the varied-cueing condition restrictively in the matched trials where participants could utilize the rejection template. Overall, we suggest that the post-stimulus mid-frontal theta also reflected the control processes of distractor filtering and implementation of rejection template in visual search.

Capitalized on color decoding of EEG signal, we demonstrated selective representations of the to-be-ignored feature in both the fixed-cueing and varied-cueing conditions. In Reeder

and colleagues' [23] fMRI study, they performed representational similarity analysis to calculate the distance metric of five cued colors in target, distractor, and neutral cueing conditions. Compared with the neutral condition, a more distinctive representation was observed only when the color was cued to be the target but not when the color was a distracting feature. The authors argued that the representations of general task-irrelevant features in neutral condition might not be an ideal baseline because participants might inhibit the neutral cue as well. Instead of comparing the representation distinctiveness across different cueing contingencies, we emphasized selective representation of the to-be-ignored feature within each condition. Meanwhile, the decoding accuracy time series provided us with a temporally fine-scaled picture of the representations of rejection templates. In another study, by repeating distractor locations, van Moorselaar and colleagues [29] showed how learned expectations promoted distractor inhibition. Whereas distractor expectation reduced distractor-specific processing after search array onset, meaningful preparatory neural tuning to distractor locations was not observed in their study. Herein, we first provide the evidence of color-selective representation of the distractor cue during the retention interval, particularly in the fixed-cueing condition. Discrepancies in the contents of rejection template as well as the experimental design might contribute to the different findings. Compared to the location-based expectation in van Moorselaar's study which only excluded one item, taking the advantage of the to-be-ignored color in the present study would help to filter out a half of the items in a search array. Moreover, fixed cues that were presented within one block could further facilitate the generation of stable rejection templates.

We found better color decoding performance in the fixed-cueing condition compared to the varied-cueing condition right after cue onset. The constant cueing contingency in fixed-cueing condition promoted the expectation of distractor color and therefore boosted the color

selectivity during the retention interval. This might account for the very early (<100 ms) emergence of the template information extracted by decoding. In fact, previous work showed that expectation could pre-activate representations of expected sensory stimuli [48] and strategically bias gazes toward the foreknown distractor in the preparatory stage to reduce stimulus-evoked attractions [49]. Behaviorally, fixed cues also facilitate distractor suppression which echoed previous findings that participants can be more effective in excluding the irrelevant items in a predictive context [14,17,20].

Given the current design, we could not elucidate how the matched distractors were filtered out based on the rejection template. Previous work found that the target template was pre-activated briefly before stimulus onset and served as a match filter to convert the sensory input of stimulus into the decision-relevant representation [50]. However, it remains unknown whether the rejection template functions the same way as the target template. It has been proposed that the rejection template might reduce the activation of matched features and therefore prevent those matched distractors from interfering with target processing in the first place [20,51]. Spatial tuning to the distractor-location also decreased when participants learned to ignore the distractor [29]. Our behavioral results showed that inhibition was achieved even in the fastest RT bin in the fixed-cueing condition, indicating a proactive suppression presumably through the reduction of the distractor-evoked response. On the contrary, rejection templates could be utilized reactively [52] which fitted our finding that suppression benefits emerged only at the longest RT bin in the varied-cueing condition. Future studies are required to further investigate the factors that modulate the implementation process.

In conclusion, our results revealed the neural markers for the distinctive processes of template formation and implementation between the fixed-cueing and varied-cueing conditions. Distractor consistency in the fixed-cueing condition promoted the formation of

rejection template and corresponded to the significant behavioral suppression benefit. In the varied-cueing condition, transient decoding performance as well as the stronger posterior alpha lateralization and mid-frontal theta/beta power during the retention interval implied the cognitive costs in template formation caused by the trial-by-trial changes of the distractor features. Differences in template formation could contribute to the utilization phase where we observed successful implementation of the rejection template under the fixed-cueing condition. This behavioral advantage of fixed-cueing was accompanied by the stronger mid-frontal theta power after search array onset in the matched trials. Together, these findings provide an integrative view of template-guided distractor inhibition in visual search.

Author contributions

WW and SL designed the research. WW, ZH and YH collected data. WW analysed data and interpreted findings under the supervision of SL. WW drafted the manuscript. SL revised the manuscript. All authors have reviewed the content and approved the final version for publication.

References

1. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
2. Soto, D., Heinke, D., Humphreys, G. W. & Blanco, M. J. Early, involuntary top-down guidance of attention from working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **31**, 248–261 (2005).
3. Soto, D., Hodsoll, J., Rotshtein, P. & Humphreys, G. W. Automatic guidance of attention from working memory. *Trends Cogn. Sci.* **12**, 342–348 (2008).
4. Reeder, R. R. & Peelen, M. V. The contents of the search template for category-level search in natural scenes. *J. Vis.* **13**, 13 (2013).
5. Eimer, M. The neural basis of attentional control in visual search. *Trends Cogn. Sci.* **18**, 526–535 (2014).
6. Stokes, M., Thompson, R., Nobre, A. C. & Duncan, J. Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19569–19574 (2009).
7. Vogel, E. K., McCollough, A. W. & Machizawa, M. G. Neural measures reveal individual differences in controlling access to working memory. *Nature* **438**, 500–503 (2005).
8. Arita, J. T., Carlisle, N. B. & Woodman, G. F. Templates for rejection: configuring attention to ignore task-irrelevant features. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 580 (2012).
9. Woodman, G. F. & Luck, S. J. Do the Contents of Visual Working Memory Automatically Influence Attentional Selection During Visual Search? *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 363–377 (2007).

10. 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).

11. Conci, M., Deichsel, C., Müller, H. J. & Töllner, T. Feature guidance by negative attentional templates depends on search difficulty. *Vis. cogn.* **0**, 1–10 (2019).

12. Han, S. W. & Kim, M. S. Do the Contents of Working Memory Capture Attention? Yes, But Cognitive Control Matters. *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 1292–1302 (2009).

13. Tanda, T. & Kawahara, J. I. Association between cue lead time and template-for-rejection effect. *Atten. Percept. Psychophys.* (2019).

14. Töllner, T., Conci, M. & Müller, H. J. Predictive distractor context facilitates attentional selection of high, but not intermediate and low, salience targets. *Hum. Brain Mapp.* **36**, 935–944 (2015).

15. Olivers, C. N. L. What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 1275 (2009).

16. Zhang, Z., Gapelin, N. & Carlisle, N. B. Probing early attention following negative and positive templates. *Attention, Perception, Psychophys.* **82**, 1166–1175 (2020).

17. Cunningham, C. A. & Egeth, H. E. Taming the White Bear: Initial Costs and Eventual Benefits of Distractor Inhibition. *Psychol. Sci.* **27**, 476–485 (2016).

18. Wen, W., Hou, Y. & Li, S. Memory guidance in distractor suppression is governed by the availability of cognitive control. *Attention, Perception, Psychophys.* **80**, 1157–1168 (2018).

19. Vatterott, D. B. & Vecera, S. P. Experience-dependent attentional tuning of distractor

rejection. *Psychon. Bull. Rev.* **19**, 871–878 (2012).

20. Noonan, M. A. P. *et al.* Distinct mechanisms for distractor suppression and target facilitation. *J. Neurosci.* **36**, 1797–1807 (2016).

21. Carlisle, N. B. Flexibility in attentional control: Multiple sources and suppression. *Yale J. Biol. Med.* **92**, 103–113 (2019).

22. Reeder, R. R., Olivers, C. N. L. & Pollmann, S. Cortical evidence for negative search templates. *Vis. cogn.* **25**, 278–290 (2017).

23. Reeder, R. R., Olivers, C. N. L., Hanke, M. & Pollmann, S. No evidence for enhanced distractor template representation in early visual cortex. *Cortex* **108**, 279–282 (2018).

24. Stilwell, B. T., Bahle, B. & Vecera, S. P. Feature-based statistical regularities of distractors modulate attentional capture. *J. Exp. Psychol. Hum. Percept. Perform.* **45**, 419–433 (2019).

25. Wang, B. & Theeuwes, J. Statistical regularities modulate attentional capture. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 13–17 (2018).

26. Ferrante, O. *et al.* Altering spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex* **102**, 67–95 (2018).

27. 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).

28. Geng, J. J., Won, B.-Y. Y. & Carlisle, N. B. Distractor Ignoring: Strategies, Learning, and Passive Filtering. *Curr. Dir. Psychol. Sci.* **28**, 096372141986709 (2019).

29. van Moorselaar, D. & Slagter, H. A. Learning What Is Irrelevant or Relevant: Expectations Facilitate Distractor Inhibition and Target Facilitation through Distinct Neural Mechanisms. *J. Neurosci.* **39**, 6953–6967 (2019).

30. Riddle, J., Scimeca, J. M., Cellier, D., Dhanani, S. & D'Esposito, M. Causal Evidence

for a Role of Theta and Alpha Oscillations in the Control of Working Memory. *Curr. Biol.* **30**, 1748-1754.e4 (2020).

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

32. Cavanagh, J. F. & Frank, M. J. Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* **18**, 414–421 (2014).

33. Cohen, M. X. A neural microcircuit for cognitive conflict detection and signaling. *Trends Neurosci.* **37**, 480–490 (2014).

34. de Vries, I. E. J., Savran, E., van Driel, J. & Olivers, C. N. L. Oscillatory Mechanisms of Preparing for Visual Distraction. *J. Cogn. Neurosci.* **31**, 1873–1894 (2019).

35. Dugué, L., Marque, P. & VanRullen, R. Theta Oscillations Modulate Attentional Search Performance Periodically. *J. Cogn. Neurosci.* **27**, 945–958 (2014).

36. Kiyonaga, A., Egner, T. & Soto, D. Cognitive control over working memory biases of selection. *Psychon. Bull. Rev.* **19**, 639–646 (2012).

37. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).

38. 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).

39. Wolff, M. J., Jochim, J., Akyürek, E. G. & Stokes, M. G. Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* **20**, 864–871 (2017).

40. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG-and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).

41. Cohen, M. X. *Analyzing Neural Time Series Data*. *Analyzing Neural Time Series Data* (MIT press, 2019). doi:10.7551/mitpress/9609.001.0001

42. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, (2011).

43. Worden, M. S., Foxe, J. J., Wang, N. & Simpson, G. V. Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J. Neurosci.* **20**, RC63 (2000).

44. Kelly, S. P., Lalor, E. C., Reilly, R. B. & Foxe, J. J. Increases in Alpha Oscillatory Power Reflect an Active Retinotopic Mechanism for Distracter Suppression During Sustained Visuospatial Attention. *J. Neurophysiol.* **95**, 3844–3851 (2006).

45. Jensen, O. & Mazaheri, A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* **4**, 186 (2010).

46. Cooper, P. S. *et al.* Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *Neuroimage* **189**, 130–140 (2019).

47. Spitzer, B. & Haegens, S. Beyond the status quo: A role for beta oscillations in endogenous content (RE)activation. *eNeuro* **4**, (2017).

48. Kok, P., Mostert, P. & de Lange, F. P. Prior expectations induce prestimulus sensory templates. *Proc. Natl. Acad. Sci.* **114**, 10473–10478 (2017).

49. Wen, W., Zhang, Y. & Li, S. Gaze dynamics of feature-based distractor inhibition under the prior-knowledge and expectation. *Atten. Percept. Psychophys.* (2021). doi:10.1101/2020.07.19.210781

50. Myers, N. E. *et al.* Testing sensory evidence against mnemonic templates. *Elife* **4**, 1–

25 (2015).

51. Moher, J., Lakshmanan, B. M., Egeth, H. E. & Ewen, J. B. Inhibition Drives Early Feature-Based Attention. *Psychol. Sci.* **25**, 315–324 (2014).

52. Moher, J. & Egeth, H. E. The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Attention, Perception, Psychophys.* **74**, 1590–1605 (2012).

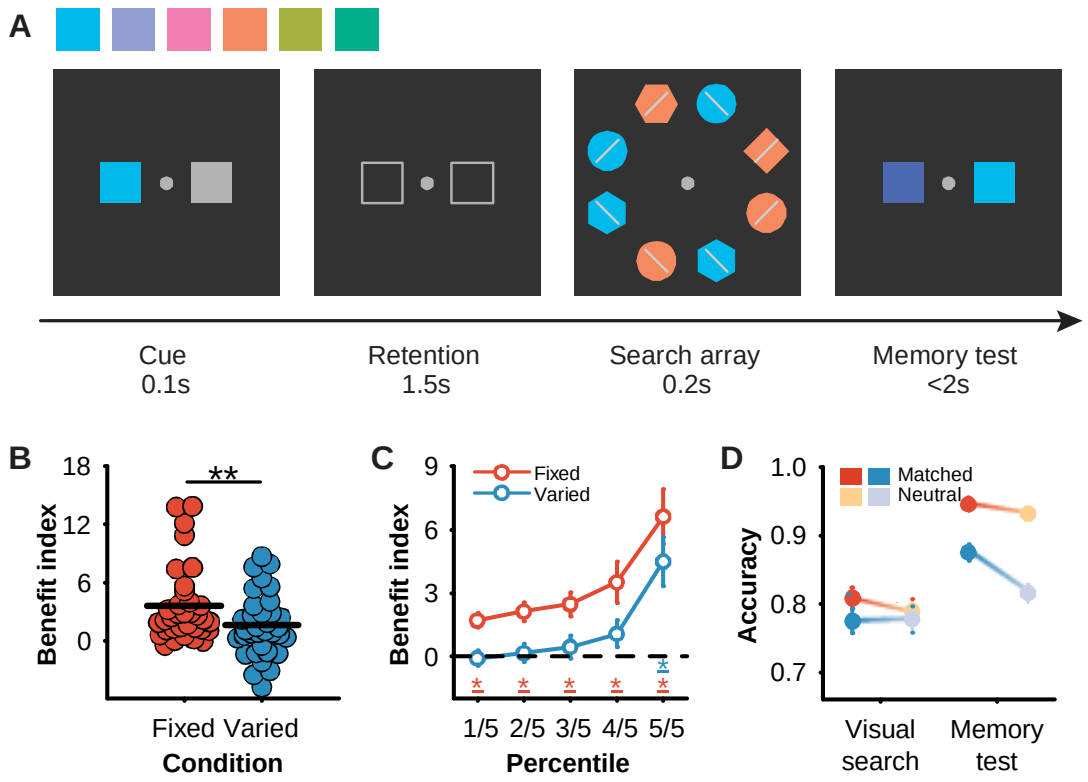

Figure 1. Task and behavioral results. (A) Task. The color cue indicated the to-be-ignored color in the search array. Participants were asked to find the target (diamond) and judge whether the bar inside of it was clockwise or counter-clockwise tilted relative to the mid-vertical. In the memory test, participants were instructed to select the color patch that matched the cue. (B) Suppression benefit index. Red and blue dots represent the individuals' benefit indices in the fixed-cueing and varied-cueing conditions, respectively. Black lines represent the averaged benefit indices across participants. (C) Suppression benefit indices of RT percentile bins. Errorbars represent the standard deviations of 5000 times bootstrapped means. (D) Accuracy of visual search task and memory test. * $p < .05$, ** $p < .01$, *** $p < .001$, $\underline{*}$ p <.001 (two-tailed).

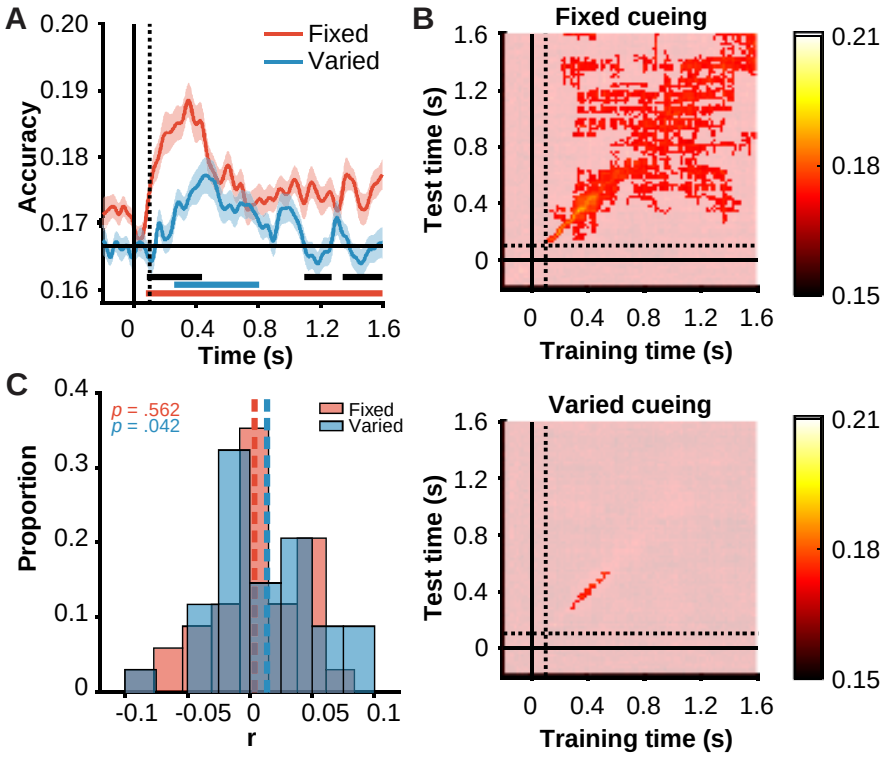

Figure 2. Results of color-decoding during retention interval. (A) Time-series of decoding accuracy in the fixed-cueing and varied-cueing conditions. The shaded areas represent the standard deviation of 5000 times bootstrapped mean. The solid/dashed vertical line indicate the cue onset/offset. Solid lines below the x-axis indicated significant time-clusters for each condition and the black lines correspond to the significant time-clusters for condition difference. (B) Cross-temporal decoding of cued color in the two cueing conditions. Alpha-blending was performed on insignificant time-points for better illustration. (C) Histogram of participants' correlation coefficients between single-trial's averaged accuracy and search RT in the matched trials. Vertical dashed lines are the average correlation coefficients across participants for the two cueing conditions. The x-axis represents the trialwise Pearson's correlation coefficients. The y-axis represents the proportion of participants. P-values are one-sample bootstrapping t-test of populations' correlation coefficients against zero.

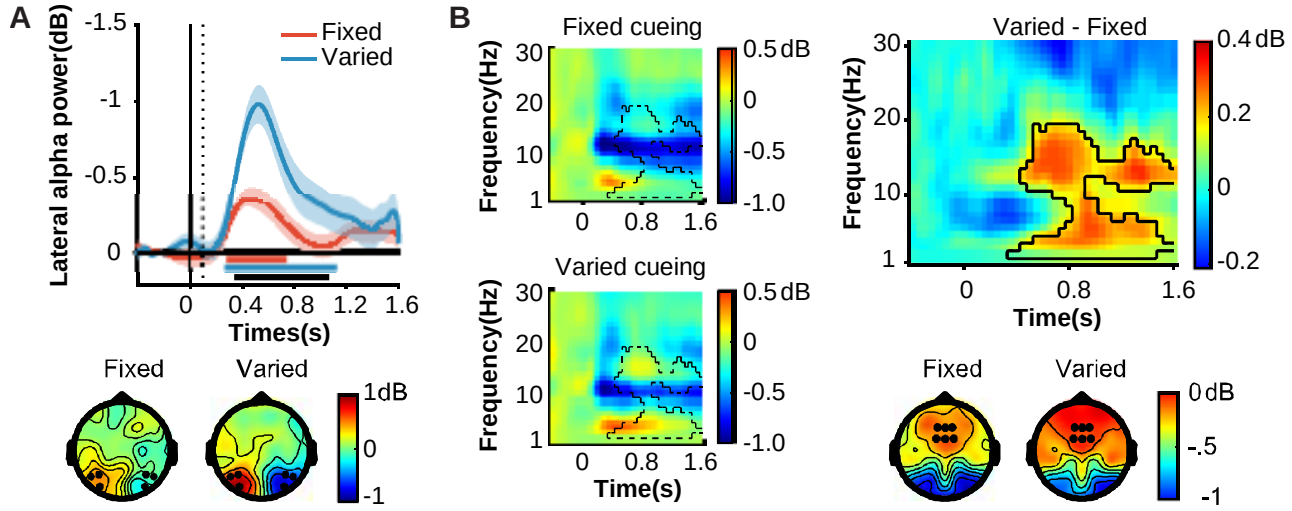

Figure 3. Oscillatory activities during retention interval. (A) Top: Posterior lateralized alpha power in the fixed-cueing and varied-cueing conditions. The shaded areas represent the between-subject standard error. The solid/dashed vertical line indicated the cue onset/offset. Solid lines below the x-axis indicated significant time-clusters for each condition with the black line corresponding to their difference. Bottom: Topography of the averaged lateralized alpha power during 300~800 ms. Topographical maps indicate the difference values of alpha power between left-cue minus right-cue trials. Since contralateral alpha power would decrease relative to ipsilateral alpha power when distractor cues are attended, left posterior electrodes are positive in this plot. The black dots highlight the channels used for alpha lateralization analysis. Black dots correspond to the channels used for alpha lateralization analysis (P5/6, P7/8, PO7/8). (B) Left: Time-frequency map extracted from mid-frontal electrodes (FCz, FC1, FC2, Fz, F1, F2). Dotted contours were adopted from the time-frequency difference map on the right column. See Supplementary Materials Figure S3 for the topography of alpha-band and beta-band activity during retention interval. Right-top: Time-frequency difference map between the varied-cueing and fixed-cueing conditions (varied minus fixed). Significant time-frequency cluster from the non-parametric permutation

test are out-lined by the dark contour. Data was extracted from mid-frontal electrodes. Right-bottom: Topographical distribution of mid-frontal theta activity (3~7 Hz, 0.1~1.6 s) for the fixed-cueing and varied-cueing conditions. Black dots correspond to the channels used for time-frequency analysis.

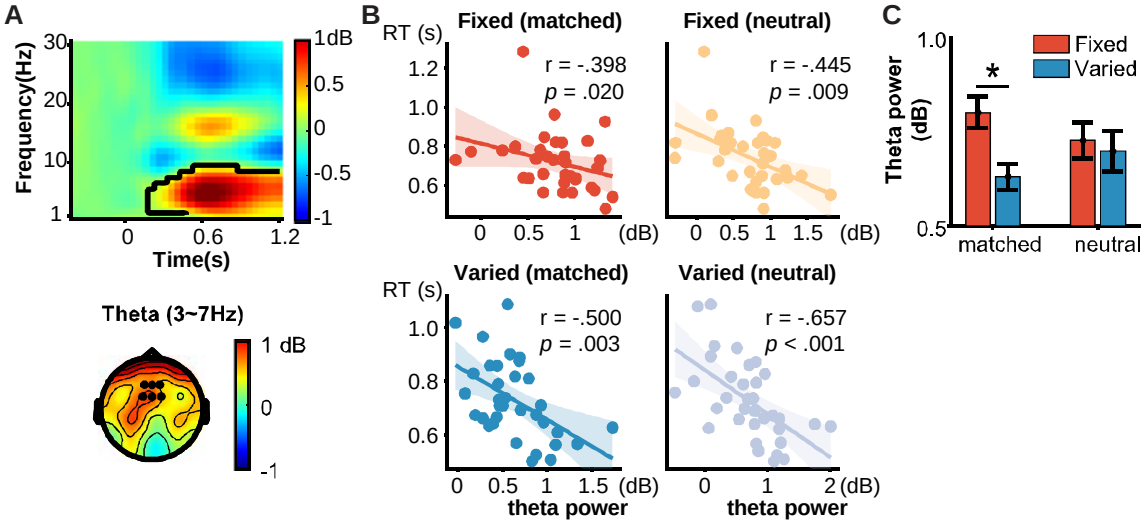

Figure 4. Mid-frontal theta after search array onset. (A) Top: Time-frequency map after search array onset. The significant time-frequency cluster from the non-parametric permutation test was outlined by the dark contour. Bottom: Topography of averaged theta power during 0.2~0.7s and the power difference between conditions. To avoid stimulus-driven and response-related contaminates, we chose the time window to be 0.2~0.7s. The averaged search RTs across conditions was 735 ms. The black dots indicated the channels used for power analysis (FCz, FC1, FC2, Fz, F1, F2). (B) Spearman-rank correlations between frontal theta power and RTs in matched trials and neutral trials of two cueing conditions. The shaded area represents the 95% confidence interval for the fitted line. (C) Mid-frontal theta power of different trial types and cueing conditions. Errorbars represent the within-subject corrected standard error.