# Visual search training benefits from the integrative effect of enhanced covert attention and optimized overt eye movements

Qi Zhang[1], Zhibang Huang[2,3,4,5], Liang Li[6], and Sheng Li[2,3,4,5]

[1] School of Educational Science, Minnan Normal University

[2] School of Psychological and Cognitive Sciences, Peking University.

[3] Beijing Key Laboratory of Behavior and Mental Health, Peking University.

[4] PKU-IDG/McGovern Institute for Brain Research, Peking University.

[5] Key Laboratory of Machine Perception (Ministry of Education), Peking University.

[6] The Brain Science Center, Beijing Institute of Basic Medical Sciences.

Author Note:

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

There are no conflicts of interest to declare.

**Abstract**

Visual search in a complex environment requires efficient discrimination between target and distractors. Training serves as an effective approach to improve visual search performance when the target does not automatically pop out from the distractors. In the present study, we trained subjects on a conjunction visual search task and examined the training effects in behavior and eye movement from Experiments 1 to 4. The results showed that training improved behavioral performance and reduced the number of saccades and overall scanning time. Training also increased the search initiation time before the first saccade and the proportion of trials in which the subjects correctly identified the target without any saccade, but these effects were modulated by stimulus' parameters. In Experiment 5, we replicated these training effects when eye movements and EEG signals were recorded simultaneously. The results revealed significant N2pc components after the stimulus onset (i.e., stimulus-locked) and before the first saccade (i.e., saccade-locked) when the search target was the trained one. These N2pc components can be considered as the neural signatures for the training-induced boost of covert attention to the trained target. The enhanced covert attention led to a beneficial tradeoff between search initiation time and the number of saccades as a small amount of increase in search initiation time could result in a larger reduction in scanning time. These findings suggest that the enhanced covert attention to target and optimized overt eye movements are coordinated together to facilitate visual search training.

# Visual search training benefits from the integrative effect of enhanced covert attention and optimized overt eye movements

Effective identification of a search target in a complex visual environment relies on multiple factors. To reveal the complex mechanisms underlying this simple behavior, investigators have developed a variety of visual search tasks in which subjects look for a pre-defined target among various distractors (Duncan & Humphrey, 1989; Theeuwes, 1992; Treisman & Gelade, 1980). According to the patterns of searching behavior, a search task is called parallel if a single feature that could easily pop-out from homogeneous distractors defines the target. As a result, reaction time (RT) in parallel search is independent of the number of distractors (Bergen & Julesz, 1983; Duncan, 1989; Egeth et al., 1972). In contrast, in a conjunction visual search task in which the target is defined by combination of two features and may share one of the two features with distractors, the task is serial as it requires shifts of attentional focus among search items (Treisman & Gelade, 1980). Consequently, reaction time in serial visual search increases as the number of distractors increase (Treisman & Paterson, 1984). Compared with parallel search, serial visual search is generally higher in task difficulty and closer to our daily visual experiences.

Our visual system evolved as a result of being optimized for intrinsic regularities of natural environment (Simoncelli & Olshausen, 2001). However, we also live in a diverse and dynamic world where experience-dependent adjustments of visual functions are necessary. Therefore, training on visual search tasks could serve as a useful probing tool for the experience-dependent visual behavior. In fact, a various forms of visual search paradigms had been adopted in visual search training studies, including visual search of shape (Qu et al., 2017; Sigman & Gilbert, 2000), orientation (An et al., 2012),

and letter (Bueichekú et al., 2016, 2019), as well as various conjunction visual search tasks (Su et al., 2014; Frank et al., 2016; Reavis et al., 2016, 2018). These studies found that training can substantially improve visual search performance. The improvement can be attributed to the increase in target-evoked activity relative to distractor-evoked activity (Reavis et al., 2016) or the transformation from serial search to parallel search (Ellison & Walsh, 1998; Sirenteanu & Rettenbach, 2000; Sireteanu & Rettenbach, 1995) as supported by the attention capture effect of trained target stimulus that elicited N2 posterior contralateral (N2pc) ERP component (Qu et al., 2017).

Visual search is a process that is generally recruits both cover attention and over attention (Hoffman & Subramaniam, 1995; Li et al., 2021; Moore & Fallah, 2001; Posner, 1980). The above-mentioned studies had adopted paradigms in which subjects were required to maintain fixation during the search task and therefore focused more on the covert aspect of attentional processing. A less investigated aspect of this line of research was the training effects on eye movement (i.e., overt attention and related processes) in conjunction visual search. In natural viewing conditions, the optimal efficiency for identifying a search target could be achieved not only by enhanced covert attentional processing and but also through improved eye movements. In the present study, we used a free viewing paradigm and recorded subjects' eye movements with an eye tracker. Therefore, we were able to investigate the changes of eye movement patterns after visual search training. In the free viewing paradigm in which the subjects can shift their gazes among search items, the duration of eye movement can be decomposed into three epochs (Malcolm & Henderson, 2009): search initiation time as the fixational period before the first saccade, scanning time as the period from the first saccade until the beginning of the last fixation, and verification time as the duration of the last fixation until the participants' response. We took the advantage of this approach and examined whether the specific

epochs and the related measurements of eye movement were modulated by visual search training.

Further, by simultaneously recording eye movements and EEG signals, we could also reveal the neural signatures of cover attention that were accompanied with the training effects in overt eye movements. Particularly, we examined saccade-locked N2pc component (Weaver et al., 2017; Talcott & Gaspelin, 2021) along with the classical stimulus-locked N2pc that measured the stimulus-induced attentional bias (Eimer, 2014; Luck & Hillyard, 1994; Hickey et al., 2006). New insights into the training-induced interactive effect of covert attention and overt eye movements could be gained given this combinatorial approach.

## Overview of the Study

We conducted four behavioral experiments (Experiments 1 to 4) and an EEG experiment (Experiment 5). Eye movement data were recorded in all experiments. Search arrays in the experiments were composed of line segments with two defining features (i.e., color and orientation, Figure 1A-C). In Experiments 1 to 3, we trained subjects in a conjunction visual search task and tested their performance before and after the training. In Experiment 4, the training was replaced with a two-hour rest period to establish a baseline performance without training. The results showed that the subjects were faster and more accurate after training. The behavioral improvement was accompanied with reduced fixation number and hence scanning time (Experiments 1 to 3), as well as increased proportion of zero-saccade correct trials and search initiation time (Experiment 1). In Experiment 5, we replicated these behavioral and eye movement findings and further identified stimulus-locked and saccade-locked N2pc components as the neural signatures for the training-induced enhancement in covert attention to the trained target. These results suggest that training on visual search increased the attentional priority

of the trained stimuli and optimized overt eye movement patterns.

We totally recruited 80 participants with 15 participants for Experiments 1 to 4 and 20 participants for Experiment 5. We conducted priori power analyses with G*Power 3.0 (Faul et al., 2007). The analyses suggested that 15 participants are required for each experiment to detect a behavioral training effect on visual search (*F* tests: repeated-measures ANOVA, within factor, effect size $\eta_p^2$ = 0.298, power = 0.9). The expected effect size was chosen based on previous studies that adopted similar conjunction visual search training paradigm (Su et al., 2014, Experiment 3a). The experiments were approved by the local ethics committee. The participants provided written informed consent prior to the experiment and were paid for their participation.

## Experiment 1

In Experiment 1, we trained participants on a conjunction visual search task. We examined the training effects on the trained target and two stimuli that either shared color or orientation feature with the trained target when they served as target in the pretest and posttest sessions.

### Method

**Participants.** Fifteen right-handed naïve participants (eight females, age range = 18–27 years, mean age = 22.5 years) with normal or corrected-to-normal vision were recruited for the experiment.

**Apparatus.** All stimuli were displayed on a 32-inch Display++ LED monitor (Cambridge Research Systems Ltd, Rochester, Kent, UK) with a refresh rate of 120 Hz and spatial resolution of 1920 × 1080 pixels. The stimuli were presented using Psychtoolbox 3.0 (Brainard, 1997; Pelli, 1997) in MATLAB programming environment (MathWorks, Natick, MA, USA). The participants were

positioned 72 cm from the monitor in a dimly lit room. We used a chin rest to fix the participants' head position. Eye movements were recorded using an EyeLink 1000 plus (SR Research, Ontario, Canada) eye tracker with a sampling rate of 1000 Hz. Gaze position was established using nine-point calibration and validation procedure. Drift correction was performed before each block.

**Stimuli.** Each search array (16.37° × 16.37°) consisted of a central fixation cross and 36 items (1.37° × 0.55°) in a 6 × 6 array (Figure 1A). The items were line segments on a gray background with distinctive colors (red or green) and orientations (0°: horizontal, 90°: vertical, 45°: tilted right from vertical, or 135°: tilted left from vertical). In half of the trials, there was a unique target (i.e., red 90°, denoted as r_90) in the search array and the rest of the items were distractors (randomly chosen from g_0, g_90, r_45, and r_135). In the other half of the trials, there was no target and all items in the search array were distractors.

**Procedure.** The procedure of the experiment is shown in Figure 1D. Participants completed a practice block (100 trials) to familiar with the task before the pretest session. In the practice block, search arrays consisted of black and white line segments with a black 90° line segment serving as target. In the pretest, participants were tested for three different targets (r_90, r_0, and g_90 with equal number of trials) separately in three blocks with 100 trials in each block. After the pretest, participants conducted eight training blocks with 400 trials in each block. The targets were always r_90 in the training blocks and appeared in a half of the trials. There was a three minutes break after every block and a five minutes break after every two blocks. During the practice and training, participants were offered an optional break every 20 trials to reduce fatigue and boredom. After training, participants completed a posttest that was identical to the pretest.

A typical trial of the conjunction visual search task is shown in Figure 1A. Each trial began with

the presentation of a central fixation cross for 400 ms, followed by a search array that appeared until the key response or the elapse of 3000 ms since its onset. Participants were instructed to fixate on the central fixation before the presentation of the search array and return to the fixation cross as soon as possible after they found the target. They could make eye movements during search period. Participants were required to indicate the presence or absence of the target as accurately and quickly as possible by pressing one of two keys with two fingers of the same hand.

**Data analysis.**

*Behavior.* Trials with incorrect responses (4.01%), RTs that were faster than 200 ms (0.09%), and latency of first saccade smaller than 80 ms (4.27%), were excluded from analyses. In the calculation of d', when participants scored perfect performance (i.e., hit rate = 1 or false alarm rate = 0), hit rate was assigned a value of 1-1/ (2 × N) or false alarm rate was assigned a value of 1/ (2 × N) (N was the total number of target-present or target-absent trials). For the presentation purpose, we refer the trials with the target of r_90, r_0, and g_90 as C+O+, C+O–, and C–O+, respectively (C for color feature, O for orientation feature, + and – indicated that the feature was the same as and different from the trained target, respectively).

Repeated-measures two-way analyses of variance (ANOVAs) with session (pretest and posttest) and target type (C+O+, C+O–, and C–O+) as factors were conducted with SPSS (version 20, IBM), and if necessary, the measures were corrected for violation of sphericity assumption by using Greenhouse–Geisser correction. Corrections made for multiple comparisons were also reported. We defined transfer index (TI) from C+O+ to either C+O– or C–O+ as the mean percent improvement (MPI) of the untrained target (C+O–, C–O+) divided by the MPI of the trained target (C+O+). The MPI of each participant and each condition was calculated as (posttest–pretest)/pretest.

*Eye movement.* We divided the correct trials into three categories: zero-saccade trials (no saccade before response), single on-target-saccade trials (only one saccade before response and this saccade was on the target), and other trials. For these other trials with over eye movement, we defined three epochs for each trial (Malcolm & Henderson, 2009): search initiation time, scanning time, and verification time. Search initiation time was defined as the period from the onset of the search display until the first saccade away from the central fixation. We assumed that search initiation time reflected the time needed to select the first item in the search display for examination. Scanning time was defined as the period from the first saccade until the onset of the last fixation. This epoch reflected the actual search process. We also calculated the number of fixations and mean fixation duration the scanning period. Verification time was defined as the duration of the last fixation until the participants' response for those target-present trials in which the last fixation was on the target. This epoch reflected the time to determine whether the fixated item was the target. Segmenting total trial duration into three epochs helped us to elucidate the training effect on these specific search processes.

Two issues need to be clarified for the definition of eye movements. First, a fixation was defined as on target or central cross if its distance from the center of the target or central cross was shorter than 1.5°. Second, we observed that the participants occasionally returned their gaze back to the central fixation to prepare for next trial but this was done before their response to the current trial. Therefore, if there were more than one fixation in a trial, and the last fixation was on the central cross and its start time was within 300 ms from the key response, we removed this last fixation from the analysis and subtracted one from the number of fixations.

The statistical analyses (i.e., ANOVAs) for the measurements of eye movement were conducted in the same way as for the behavioral measurements.

## Results

**Behavior.** The behavioral results and statistics of Experiment 1 are shown in Figure 2A and Table 1A, respectively. For both d' and RTs, we observed significant effects on session and target type, but not their interaction. These results showed that training improved visual search performance.

To examine the transfer of learning effect to the stimuli that differed from the trained target in one of the two features, we calculated the transfer indices for C+O– (d': 0.91±1.34; RTs: 0.95±0.39) and C–O+ (d': 0.02±1.88; RTs: 0.74±0.35) conditions. Paired t-tests revealed no significant difference in transfer index between C+O– and C–O+ conditions in d' ($t(14) = 1.81$, $p = 0.09$, Cohen's $d = 0.47$) and in RTs ($t(14) = 1.81$, $p = 0.09$, Cohen's $d = 0.47$).

**Eye movement.** The eye movement results and statistics of Experiment 1 are shown in Figure 2B and Table 1B, respectively. The trials were grouped as zero-saccade trials, single on-target-saccade trials, and other correctly responded trials.

The proportions of the zero-saccade trials and single on-target-saccade trials increased significantly after training. For other correctly responded trials, training significantly increased the search initiation time and reduced the number of fixations and scanning time. However, the mean fixation duration or verification time was not significantly modulated by training for these other trials.

## Discussion

The behavioral results in Experiment 1 demonstrated that training on the conjunction visual search task significantly improved search performance of the participants as they identified the target faster and more accurate. This improvement was not specific to the trained target stimulus, but also transferred to the stimuli that shared one of the two features with the trained target (i.e., C+O– and C–

O+), as was evident with the significant session effect and the absence of its interaction with target type. The eye movement results revealed that the reduced RTs after training was mainly the result of the reduction in the number of fixations (or saccades). In contrary, mean fixation duration and verification time remained unchanged after training, indicating that the processing time for each fixated item was not affected by training.

There were two interesting findings in the eye movement results. First, we found that the proportions of the zero-saccade trials and single on-target-saccade trials significantly increased after training, suggesting that the participants had a higher chance to identify the target before making any saccade. Second, we observed an increase in search initiation time after training for other correctly responded trials. This increase might reflect a change in search strategy as it could increase the probability of correct identification of target with less saccades. The tradeoff between search initiation time and the number of saccades was beneficial as a small amount of increase in search initiation time could lead to a larger reduction in scanning time. The potential mechanism that could account for these two findings was that training improved the covert attention to the trained stimulus. We will examine this interpretation in Experiment 5 where EEG signals were recorded.

## Experiment 2

In Experiment 2, we used a search array with a larger number of items that were smaller in size, to examine whether the observed training effects from Experiment 1 could be generalized to a different set of stimulus parameters. The overall area of the search array remained unchanged to keep the attentional span comparable across experiments.

## Method

**Participants.** Fifteen right-handed naïve participants (12 females, age range = 19–28 years, mean age = 23.7 years) with normal or corrected-to-normal vision were recruited for the experiment.

**Apparatus.** The apparatus in Experiment 2 was identical to Experiment 1.

**Stimuli.** The stimuli in Experiment 2 were identical to Experiment 1, except that each search array consisted of 144 items (12 ×12) which were smaller in size (0.68° × 0.27°) (Figure 1B).

**Procedure.** The general procedure in Experiment 2 was identical to Experiment 1. As the task would become more difficult with the increased set size and smaller search items, participants completed two practice blocks with 100 trials in each block to familiar with the task before the pretest. In the pretest session, we tested three targets (r_90, r_180, and g_90) with 160 trials per target to get enough correct trials for the analysis. Target was always r_90 in eight training blocks with 300 trials in each block, keeping similar training time with Experiment 1.

**Data analysis.** The data analysis was identical to Experiment 1. Trials with incorrect responses (24.97%), RTs that were faster than 200 ms (0.12%), and latency of first saccade smaller than 80 ms (7.85%), were excluded from analyses.

## Results

**Behavior.** The behavioral results and statistics of Experiment 2 are shown in Figure 3A and Table 2A, respectively. For d', we observed significant effects on session, target type, and their interaction. Simple-effects analysis (Bonferroni-corrected) revealed significant effects of training in C+O+ ($t(14) = -5.91$, $p < 0.001$, Cohen's $d = 1.53$), C+O– ($t(14) = -4.18$, $p < 0.001$, Cohen's $d = 1.08$), and C–O+ ($t(14) = -3.66$, $p < 0.01$, Cohen's $d = 0.94$) conditions. For RTs, we also observed significant

effects on session, target type, and their interaction. Simple-effects analysis (Bonferroni-corrected) revealed significant effects of training in C+O+ ($t(14) = 9.27$, $p < 0.001$, Cohen's $d = 2.39$), C+O– ($t(14) = 4.57$, $p < 0.001$, Cohen's $d = 1.18$), and C–O+ ($t(14) = 3.52$, $p < 0.01$, Cohen's $d = 0.91$) conditions.

Paired t-tests on the transfer indices for C+O– (d': 0.64±0.72; RTs: 0.74±0.48) and C–O+ (d': 0.32±0.55; RTs: 0.30±0.37) conditions revealed a significant difference in RTs ($t(14) = 2.55$, $p < 0.05$, Cohen's $d = 0.66$, TI was larger for C+O–) but not in d' ($t(14) = 1.66$, $p = 0.12$, Cohen's $d = 0.43$).

To examine the task difficulty between Experiments 1 and 2, we conducted a mixed design ANOVAs on d' and RTs of pretest session with target types (C+O+, C+O– and C–O+ conditions) and experiment (Experiment 1 and Experiment 2) as factors. The results revealed that there were significant effects on experiment (d: $F(1, 28) = 81.82$, $p < 0.001$, $\eta_p^2 = 0.75$; RTs: $F(1, 28) = 74.46$, $p < 0.001$, $\eta_p^2 = 0.73$). No significant effect was found on target types (d: $F(2, 56) = 1.48$, $p = 0.24$, $\eta_p^2 = 0.05$; RTs: $F(2, 56) = 0.95$, $p = 0.39$, $\eta_p^2 = 0.03$) or its interaction with experiment (d: $F(2, 56) = 2.21$, $p = 0.12$, $\eta_p^2 = 0.07$; RTs: $F(2, 56) = 0.91$, $p = 0.41$, $\eta_p^2 = 0.03$).

**Eye movement.** The eye movement results and statistics of Experiment 2 are shown in Figure 3B and Table 2B, respectively.

Training did not significantly increase the proportion of the zero-saccade trials and the search initiation time of other correctly responded trials in Experiment 2. However, there was still a significant increase in the proportion of the single on-target-saccade trials. Like Experiment 1, training significantly reduced the number of fixations and scanning time and kept the mean fixation duration and verification time unchanged for the other correctly responded trials. There was a significant interaction in the mean fixation duration, but simple-effects analysis (Bonferroni-corrected) revealed

no significant effect of training in C+O+ ($t(14) = 1.48$, $p = 0.16$, Cohen's $d = 0.38$), C+O– ($t(14) = -0.18$, $p = 0.86$, Cohen's $d = 0.05$), or C–O+ ($t(14) = -1.61$, $p = 0.13$, Cohen's $d = 0.42$) condition. Additionally, the training effects in the number of fixations and scanning time interacted with trial type. Simple-effects analyses (Bonferroni-corrected) revealed significant effects of training in C+O+ ($t(14) = 10.45$, $p < 0.001$, Cohen's $d = 2.70$), C+O– ($t(14) = 7.02$, $p < 0.001$, Cohen's $d = 1.81$), and C–O+ ($t(14) = 4.06$, $p < 0.01$, Cohen's $d = 1.05$) conditions for the number of fixations, and significant effects of training in C+O+ ($t(14) = 9.10$, $p < 0.001$, Cohen's $d = 2.35$), C+O– ($t(14) = 5.09$, $p < 0.001$, Cohen's $d = 1.31$) and C–O+ ($t(14) = 3.19$, $p < 0.01$, Cohen's d = 0.82) conditions for the scanning time.

**Discussion**

With the increased task difficulty, we observed similar patterns of results in Experiment 2 as those in Experiment 1, except that no significant training effects for the proportion of the zero-saccade trials and the search initiation time of other correctly responded trials were observed. Apparently, increasing the number of items in the search display and reducing the items' size made it less unlikely that the participants could locate the target when fixated at the center of the display. It was possible that the smaller item-to-item distance in Experiment 2 introduced crowding effect (Levi, 2008; Whitney & Levi, 2011) that prevented the participants properly perceived the items at their periphery visual field. To address this issue, in Experiment 3, we remained the item size unchanged as in Experiment 2 but changed the item number to be the same as in Experiment 1, making the items in the search array less crowded than both Experiment 1 and Experiment 2.

## Experiment 3

In Experiment 3, we kept the size of the search items identical to Experiment 2 but had the number of items (6 × 6) to be the same as in Experiment 1 (Figure 1C). Importantly, this manipulation resulted in a less crowded search array as compared with both Experiment 1 and Experiment 2 while the overall attentional span remained unchanged, thus allowing us to investigate whether crowding effect was the key factor that contribute the different training effects between Experiments 1 and 2.

### Method

**Participants.** Fifteen right-handed naïve participants (eight females, age range = 18–24 years, mean age = 21.1 years) with normal or corrected-to-normal vision were recruited for the experiment.

**Apparatus.** The apparatus in Experiment 3 was identical to Experiment 1.

**Stimuli.** The stimuli in Experiment 3 were identical to Experiment 2, except that each search array consisted of 36 items (6 × 6).

**Procedure.** The general procedure in Experiment 3 was identical to Experiment 2.

**Data analysis.** The data analysis was identical to Experiment 1. Trials with incorrect responses (6.00%), RTs that were faster than 200 ms (0.01%), and latency of first saccade smaller than 80 ms (3.72%), were excluded from analyses.

### Results

**Behavior.** The behavioral results and statistics of Experiment 3 are shown in Figure 4A and Table 3A, respectively. For d’, we observed significant effects on session and its interaction with target type. No significant effect on target type was observed. Simple-effects analysis (Bonferroni-corrected)

revealed significant effect of training in C+O+ ($t(14) = -7.27$, $p < 0.001$, Cohen's $d = 1.88$) condition. No significant effect of training was observed in C+O– ($t(14) = -1.47$, $p = 0.16$, Cohen's $d = 0.38$) or C–O+ ($t(14) = -1.29$, $p = 0.22$, Cohen's $d = 0.33$) condition. For RTs, we observed significant effects on session and target type, but not their interaction.

Paired t-tests on the transfer indices for C+O– (d': -0.01±1.54; RTs: 0.86±0.80) and C–O+ (d': -0.25±1.77; RTs: 0.77±0.33) conditions found significant difference neither in d' ($t(14) = 1.00$, $p = 0.33$, Cohen's $d = 0.26$) nor in RTs ($t(14) = 0.61$, $p = 0.55$, Cohen's $d = 0.16$).

To examine the task difficulty between Experiments 1 and 3, we conducted a mixed design ANOVAs on d' and RTs of pretest session with target types and experiment as factors. For d', the results revealed significant interaction between the two factors ($F(2, 56) = 6.73$, $p < 0.01$, $\eta_p^2 = 0.19$). No significant effect was found on experiment ($F(1, 28) = 0.26$, $p = 0.62$, $\eta_p^2 = 0.01$) or target types ($F(2, 56) = 0.13$, $p = 0.88$, $\eta_p^2 < 0.01$). Simple-effects analysis (Bonferroni-corrected) only revealed significant difference between C+O+ and C-O+ condition in Experiment 1 ($t(14) = 2.86$, $p < 0.05$, Cohen's $d = 0.74$). For RTs, we observed significant effect on target types ($F(2, 56) = 3.64$, $p < 0.05$, $\eta_p^2 = 0.11$). No significant effect was found on experiment ($F(1, 28) = 3.43$, $p = 0.07$, $\eta_p^2 = 0.11$) or their interaction ($F(2, 56) = 0.90$, $p = 0.41$, $\eta_p^2 = 0.03$).

**Eye movement.** The eye movement results and statistics of Experiment 3 are shown in Figure 4B and Table 3B, respectively.

The eye movement results were similar to Experiment 2. We observed a significant increase in the proportion of the single on-target-saccade trials, but not in the proportion of the zero-saccade trials and the search initiation time of other correctly responded trials. Training significantly reduced the number of fixations and scanning time. There was a significant decrease in mean fixation duration.

Verification time remained unchanged after training.

**Discussion**

By reducing the size of the search items and keeping the number of items unchanged as in Experiment 1, we observed similar results as in Experiment 2. These results showed that crowding effect was unlikely the key factor that determined whether training could introduce a change in the proportion of the zero-saccade trials and the search initiation time of other correctly responded trials. Unlike Experiment 2, there was no significant difference in task difficulty between Experiments 1 and 3, further excluding task difficulty as a critical factor that contributed to the differential training effects between experiments. Overall, the findings of Experiments 1 to 3 suggest that the main behavioral and eye movement effects were consistent across experiments but the stimulus' parameters could also modulate the training effects. Future investigations are required to elucidate the modulatory effects of specific parameters.

## Experiment 4

In Experiment 4, we examined the practice effect that could be induced by the test sessions. We adopted the same procedure as in Experiment 2 except that there was no training session.

**Method**

**Participants.** Fifteen right-handed naïve participants (11 females, age range = 18–24 years, mean age = 19.6 years) with normal or corrected-to-normal vision were recruited for the experiment.

**Apparatus.** The apparatus in Experiment 4 was identical to Experiment 1.

**Stimuli.** The stimuli in Experiment 4 were identical to Experiment 2 (Figure 1B).

**Procedure.** The procedure in Experiment 4 were identical to Experiment 2, except that there was no training in Experiment 4 (Figure 1E). After the pretest session, the participants had a two-hour break before the posttest session.

**Data analysis.** The data analysis was identical to Experiment 1. Trials with incorrect responses (31.78%), RTs that were faster than 200 ms (0.17%), and latency of first saccade smaller than 80 ms (6.68%), were excluded from analyses.

## Results

**Behavior.** The behavioral results and statistics of Experiment 4 are shown in Figure 5A and Table 4A, respectively. For both d' and RTs, we observed significant effects on session, but not on target type and their interaction.

We used a mixed design ANOVA to compared the improved d' and RTs between Experiment 2 and Experiment 4. We calculated the improvement of d' (posttest–pretest) and RTs (pretest–posttest) in two experiments separately. The mixed design ANOVA with target types (C+O+, C+O– and C–O+ conditions) and experiment (Experiment 2 and Experiment 4) as factors were conducted on improved d' and RTs. For d' improvement, there were significant effects on experiment ($F(1, 28) = 24.15$, $p < 0.001$, $\eta_p^2 = 0.46$), target type ($F(1.52, 42.58) = 4.18$, $p < 0.05$, $\eta_p^2 = 0.13$), and their interaction ($F(1.52, 42.58) = 10.21$, $p < 0.001$, $\eta_p^2 = 0.27$). Simple-effects analysis (Bonferroni-corrected) revealed significant effects of experiment in C+O+ ($t(18.72) = 5.18$, $p < 0.001$, Cohen's $d = 1.89$) and C+O– ($t(20.18) = 2.70$, $p < 0.05$, Cohen's $d = 0.99$) conditions. No significant effect of experiment was found in C–O+ condition ($t(23.95) = 1.15$, $p = 0.26$, Cohen's $d = 0.42$). For RTs improvement, there were

significant effects on experiment ($F(1, 28) = 18.43$, $p < 0.001$, $\eta_p^2 = 0.40$), target type ($F(2, 56) = 4.99$, $p < 0.05$, $\eta_p^2 = 0.15$), and their interaction ($F(2, 56) = 5.12$, $p < 0.01$, $\eta_p^2 = 0.15$). Simple-effects analysis (Bonferroni-corrected) revealed significant effect of experiment in C+O+ ($t(28) = 5.48$, $p < 0.001$, Cohen's $d = 2.00$) and C+O– ($t(28) = 2.51$, $p < 0.05$, Cohen's $d = 0.92$) conditions. No significant effect of experiment was found in C–O+ condition ($t(28) = 1.38$, $p = 0.18$, Cohen's $d = 0.50$).

**Eye movement.** The eye movement results and statistics of Experiment 4 are shown in Figure 5B and Table 4B, respectively.

There was no significant effect observed for the proportions of the zero-saccade trials, the proportion of single on-target-saccade trials, and the search initiation time of other correctly responded trials. For the other correctly responded trials, there were significant effects of session in the number of fixations and scanning time and a significant effect of trial type in the mean fixation duration. Significant interaction was observed between the two factors in verification time. Simple-effects analysis (Bonferroni-corrected) revealed significant difference between pretest and posttest in C+O+ ($t(14) = 2.33$, $p < 0.05$, Cohen's $d = 0.60$) condition. No significant effect was found in C+O– ($t(14) = -1.04$, $p = 0.32$, Cohen's $d = 0.27$) or C–O+ ($t(14) = 0.34$, $p = 0.74$, Cohen's $d = 0.09$) condition. No other significant effects were observed.

We used the mixed design ANOVA to compare the changes in the number of fixations and scanning time of the other correctly responded trials between Experiment 2 and Experiment 4. For number of fixations, we observed significant effects on experiment ($F(1, 28) = 34.14$, $p < 0.001$, $\eta_p^2 = 0.55$) and its interaction with trial type ($F(2, 56) = 4.60$, $p < 0.05$, $\eta_p^2 = 0.14$). Simple-effects analysis (Bonferroni-corrected) revealed significant effects of experiment in C+O+ ($t(28) = -5.61$, $p < 0.001$, Cohen's $d = 2.05$) and C+O– ($t(28) = -3.94$, $p < 0.001$, Cohen's $d = 1.44$) conditions. There was a trend

of significance in C–O+ condition ($t(28) = -1.96$, $p = 0.06$, Cohen's $d = 0.72$). For scanning time, there were significant effects on experiment ($F(1, 28) = 18.27$, $p < 0.001$, $\eta_p^2 = 0.39$), target type ($F(2, 56) = 5.10$, $p < 0.01$, $\eta_p^2 = 0.15$), and their interaction ($F(2, 56) = 4.76$, $p < 0.05$, $\eta_p^2 = 0.15$). Simple-effects analysis (Bonferroni-corrected) revealed significant effects of experiment in C+O+ ($t(28) = -5.06$, $p < 0.001$, Cohen's $d = 1.85$) and C+O– ($t(28) = -2.67$, $p < 0.05$, Cohen's $d = 0.98$) conditions. No significant effect was observed in C–O+ condition ($t(28) = -1.17$, $p = 0.25$, Cohen's $d = 0.43$).

**Discussion**

Between-experiment comparisons in behavioral and eye movement results showed that training effects (Experiment 2) were significantly larger than practice effects without training (Experiment 4). These results confirmed that our training protocols were effective despite that the whole experiment was completed within a single day.

## Experiment 5

In Experiment 5, we aimed to investigate the physiological signals that may contribute to the observed behavioral and eye movement training effects in Experiments 1, 2, and 3. We trained the participants with the same stimuli as in Experiment 1 and tested them with EEG and eye movement recorded after the training. Specifically, we examined the difference between conditions in which trained and untrained stimuli served as search target during the test. N2pc was initially identified as a key component that reflects the focusing of spatial attention onto the target location (Eimer, 2014; Luck & Hillyard, 1994). Following studies had further demonstrated that N2pc could also serve as a reliable indicator of attentional capture by physically salient stimulus (Hickey et al., 2006). In Experiment 5,

we measured N2pc for the stimuli of interest (i.e., the trained target stimulus and an untrained stimulus) when they served as target or distractor during the conjunction visual search task. Specifically, we examined the traditional stimulus-locked N2pc, as well as the presaccadic (i.e., saccade-locked) N2pc that was suggested to indicate the covert attention before overt eye movement (Weaver et al., 2017; Talcott & Gaspelin, 2021).

**Method**

**Participants.** Twenty right-handed naïve participants (nine females, age range = 18–26 years, mean age = 21.9 years) with normal or corrected-to-normal vision were recruited for the experiment.

**Apparatus.** All stimuli were displayed on a CRT monitor with a refresh rate of 60 Hz and spatial resolution of 1024 × 768 pixels. We controlled the stimulus presentation with Psychtoolbox 3.0 (Brainard, 1997; Pelli, 1997) in MATLAB programming environment (MathWorks, Natick, MA, USA). Participants were positioned 60 cm from the monitor in a dimly lit room. We used a chin rest to fix participants' head position. EEG and eye movement data were simultaneously recorded. Eye movements were recorded using an EyeLink 1000 plus (SR Research, Ontario, Canada) eye tracker with a sampling rate of 1000 Hz. Gaze position was established using nine-point calibration and validation scheme. A drift correction was carried out before each block. EEG was digitized on-line at a sampling rate of 1000 Hz and recorded from using BrainAmp amplifier (Brain Products GmbH, Gilching, Germany) from a 64-channel cap (Easycap GmbH, Germany) with 64 Ag/AgCl electrodes arranged according to the International 10-20 System. The vertical electro-oculogram (EOG) was recorded below the right eye. Electrode impedances were kept below 10 kΩ during experiment.

**Stimuli.** The stimuli in Experiment 5 were identical to Experiment 1, except that g_45 could

also serve as target during the training and test sessions.

**Procedure.** As shown in Figure 6A, each trial began with the presentation of a central fixation cross (presented for 450/500/550/600/650 ms) followed by a search array that appeared for 3000 ms or until key response. The inter-trial interval was 400/500/600 ms in practice and training sessions and 1400/1500/1600 ms in EEG test session to prevent the interference on EEG signal from the previous trial. These variations were introduced as a means of inter-trial jittering and to reduce the influence of participants' expectation. We chose longer inter-trial interval in EEG session to reduce the potential interference of EEG signals between adjacent trials.

Participants completed the experiment in two days (Figure 6B). On the first day, they completed one practice block (120 trials with black and white line segments, target was black 90° line segments) to be familiar with the task and key response before training. During training (10 blocks with 300 trials in each block), participants were provided a break every 20 trials to reduce fatigue and boredom. On the second day, they completed 10 test blocks (120 trials in each block) with EEG and eye movements recorded.

During the training session, r_90 served as trained target for a half of participants and g_45 served as trained target for other half of participants. Also, g_0, g_90, r_45, and r_135 always served as distractors with an additional distractor being the untrained target (the target in other group of participants). In each search display, a half of items were in green color (g_0 and g_90 with equal numbers) and the other half of items were in red (r_45 and r_135 with equal numbers). The untrained target randomly replaced one of the distractors in the same color.

During test session with EEG recording, the trained target was tested in first five blocks where the untrained target was always presented as distractor, and the untrained target was tested in later five

blocks where the trained target was always presented as distractor. Like the training session, there were equal number of green and red items in each search display. The untrained or trained target randomly replaced one of the original distractors in its color when it served as distractor in the block.

**Data analysis.** Analyses of behavioral and eye movement data were identical to Experiment 1, except that paired t-tests between the trained target condition and untrained target condition during the test session were used to examine the training effects. Note that whether r_90 or g_45 served as the trained target or untrained target was counterbalanced across participants, we expected that difference between the trained target and untrained target conditions in the test session was due to the training. Same trial exclusion criterion was applied in ERP analysis.

*EEG preprocessing.* EEGLAB toolbox (v14.1.2; Delorme & Makeig, 2004) and EYE-EEG extension to EEGLAB (v0.81; Dimigen et al., 2011) in MATLAB environment were used for data analysis. All EEG data were re-referenced offline to the average of the left and right mastoids (TP9 and TP10) and then filtered with a 1 Hz high pass filter and a 40 Hz low pass filter. Eye tracking data were imported and synchronized with EEG signals. After synchronization, EEGs were segmented into epochs beginning at 100 ms before stimulus onset and ending at 600 ms after stimulus onset. Epochs that contained deteriorated eye tracking data (eye blinks or lost tracking) were removed (0.91% of all trials). Epochs that contained extreme EEG amplitudes (exceeded ±50 μV) of PO7 and PO8 channels were also removed (0.12% of all trials). Independent components were extracted from the combined EEG and eye movement data. The artifactual components were rejected based on the covariance with eye movement data using saccade to fixation variance ratio criterion of 1.1 (Plöchl et al., 2012). Baseline for event related potentials (ERPs) was set to be the mean voltage of 100 ms before stimulus onset.

*Stimulus-locked N2pc.* N2pc elicited by a specific item was defined as the differential ERP between contralateral and ipsilateral posterior scalp positions with respect to the item's location. N2pc amplitudes were calculated for the four conditions defined by the corresponding item's training history (trained vs. untrained) and role in the task (target vs. distractor) using the data of PO7 and PO8 electrodes. The two electrodes were averaged together to form an overall contralateral minus ipsilateral difference wave for each condition. To determine the period in which the averaged stimulus-locked N2pc was significantly deviated from baseline, successive t-test was used with a moving window of 20 ms (20 time points) in steps of 1 ms (1 time point). The criterion of the onset time of N2pc was that at least 40 consecutive windows (40 ms) reached significance at $p < 0.05$ level (Qu et al., 2017).

*Saccade-locked N2pc.* We identified and analyzed the N2pc in ERPs time-locked to the onset of the first saccade (Weaver et al., 2017). Trials with initial saccade latencies that below 80 ms (anticipations) or above 600 ms (retardations) were discarded. For the N2pc component locked to the first saccade, a criterion of at least 20 consecutive windows (20 ms) reached significance at $p < 0.05$ level was used to determine its onset time.

## Results

**Behavior.** The behavioral results and statistics of Experiment 5 are shown in Figure 7A and Table 5A, respectively. Significant effects of session were observed for both d' and RTs.

**Eye movement.** The eye movement results and statistics of Experiment 5 are shown in Figure 7B and Table 5B, respectively.

Training significantly increased the proportions of the zero-saccade trials and single on-target-

saccade trials. For the other correctly responded trials, training significantly increased the search initiation time and reduced the number of fixations and scanning time. No significant training effect was observed in mean fixation duration or verification time for these other trials.

**EEG.** We examined the stimulus-locked and saccade-locked N2pc components for four conditions in which the trained and untrained stimuli served as target or distractor (TT: trained stimulus as target, TD: trained stimulus as distractor, UT: untrained stimulus as target, UD: untrained stimulus as distractor). The target-present trials were used in the analyses for N2pc components. Zero-saccade trials were analyzed as a separate condition for the stimulus-locked N2pc component.

For the stimulus-locked N2pc components (Figure 8A), we observed a significant N2pc when the trained stimulus served as target in the search task (TT: 234~415 ms, peaked at 342 ms). No significant N2pc was found in other three conditions (i.e., TD, UT, and UD). Significant N2pc was also observed for the zero-saccade trials in the TT condition (288~343 ms, peaked at 309 ms, 12 participants that had at least 10 zero-saccade trials in this condition were included in this analysis).

For the saccade-locked N2pc components (Figure 8B), significant N2pc was only evident when the trained stimulus served as target (TT: -94~0 ms). No significant saccade-locked N2pc was found in other three conditions.

## Discussion

The behavioral and eye movement results demonstrated significant training effects when comparing the two conditions in which the trained and untrained stimulus served as the search target, respectively. These results were comparable to the observed effects in Experiment 1.

The EEG results revealed both stimulus-locked and saccade-locked N2pc components that were

elicited by the trained stimulus when it served as target in the test session. Specifically, the untrained stimulus shared no feature with the trained target and did not elicit significant N2pc component. The significant stimulus-locked N2pc component for the trained target condition, especially for those of the zero-saccade trials, suggest an enhanced covert attention to the trained target. A recent study has shown that the presaccadic N2pc component could be observed only if covert attention was allocated to the stimulus before the generation of the eye movement (Talcott & Gaspelin, 2021). Hence, the saccade-locked N2pc could also be explained as the enhanced covert attention to the trained target. Taken together, we suggest that the increased covert attentional processing to the trained stimuli was the underlying mechanism that drove the behavioral and the overt eye movement changes after training.

## General Discussion

The present study adopted a conjunction visual search task and examined the effects of single-day training on the task performance. We measured eye movements and EEG signals to elucidate the potential mechanisms underlying the training effects. The results showed that training on the conjunction visual search task led to significant behavioral effects on d' and RTs (Experiments 1, 2, 3, and 5) and such improvement was significantly reduced when training was not taking place (Experiment 4). Meanwhile, the eye movement data revealed that the behavioral improvement was accompanied with reduced fixation number in all training protocols (Experiments 1, 2, 3, and 5) and increased proportion of zero-saccade trials and search initiation time of other correctly responded trials only in Experiments 1 and 5. This between-experiment difference could be attributed to the stimulus' parameters that varied between them. Finally, we identified stimulus-locked and saccade-locked N2pc components elicited by the trained target stimulus as the neural signatures for training-induced

enhancement in covert attention (Experiment 5). These findings collectively suggest that perceptual training on visual search increases the attentional priority of the trained stimuli and facilitates behavioral performance through optimized eye movement patterns.

In our experiments, we adopted a free viewing paradigm for the conjunction visual search task and was able to examine the patterns of eye movements in addition to the traditional behavioral measurements (Baluch & Itti, 2010). Free viewing is a natural way of observation when our visual system functions in various visual tasks. Allowing subjects to move their eyes during the search tasks was suggested as a key factor for understanding the mechanism of selective attention in complex natural environment (Hayhoe & Ballard, 2005). Importantly, we found significant behavioral improvement after training that was consistent with previous studies that required subjects keeping fixation during the task (An et al., 2012; Bueichekú et al., 2016, 2019; Clark et al., 2015; Leonards et al., 2002; Qu et al., 2017; Reavis et al., 2016, 2018; Sigman & Gilbert, 2000; Sirenteanu & Rettenbach, 2000; Sireteanu & Rettenbach, 1995; Su et al., 2014). Like previous studies, we found improved d' and reduced RTs, as well as the transfer of training effects to stimuli that shared a feature with the trained target. The transfer effect was consistent with previous investigations showing non-specific training effect of visual search task (Lee et al., 2018; Sirenteanu & Rettenbach, 2000; Sireteanu & Rettenbach, 1995; Su et al., 2014). These results suggest that the free viewing paradigm is a valid approach to investigate visual search training.

Our analyses were based on correct trials and these trials were assigned to three groups. Among them, the zero-saccade trials had no saccade during the search. In the other two groups, the single on-target-saccade trials and other correctly responded trials, at least one saccade was executed during the search period. In Experiment 5, we observed larger proportion of zero-saccade correct trials for the

trained than untrained target. This difference was accompanied by a significant stimulus-locked N2pc for the trained target, indicating stronger covert attention in this condition. For the single on-target-saccade trials and other correctly responded trials, the search initiation time was longer for the trained than untrained target. The observed difference in these saccadic trials was accompanied by a significant saccade-locked N2pc for the trained target. This presaccadic N2pc component has been suggested to reflect the covert attention before the generation of the eye movement (Talcott & Gaspelin, 2021; but see Li et al., 2021 for potential dessociation of presaccadic and covert attention). Therefore, we could conclude that training enhanced the covert attention to the trained target stimulus, at least in the zero-saccade trials. This enhancement led to either larger proportion of trials in which the participants could identify the target while fixated the central fixation or increased search initiation time. Specifically, though increasing search initiation time showed apparent detrimental effect on search speed, it nevertheless provided the opportunity to identify the target without further eye movements that could greatly save the search time. The results in Experiment 5 suggest that training enhanced covert attentional processing towards the training target stimulus and could underlie the changed overt eye movement patterns.

Previous studies found that the training effect could also be observed while the trained target served as distractors (Qu et al., 2017), an effect that was not observed in our experiment. There were two possibilities that may account for the difference between the studies. First, the tasks and stimuli were different in the two studies. We used a conjunction visual search task with color and orientation as the critical features. In Qu et al. (2017), shape served as the only critical feature in the visual search task. Second and more importantly, we adopted a short training protocol that was completed within a single day, whereas the participants in Qu et al. (2017) were trained for several days. More extensive

training is likely to result in stronger attentional bias to the trained stimulus and this elevated attentional processing could facilitate its competition against the search target for attentional selection.

In summary, with a free-viewing conjunction visual search training paradigm, we found significant improvement in behavioral performance after training. The behavioral training effects were accompanied by reduced number of saccades, as well as increased proportion of zero-saccade trials and search initiation time that were influenced by stimulus parameters. EEG results showed both stimulus-locked and saccade-locked N2pc components when the search target was the trained one, indicating training-induced enhancement in covert attention. Taken together, our findings offered new insights that the combination of the enhanced covert attention to target and optimized overt eye movements contribute to the behavioral improvement after visual search training.

## References

An, A., Sun, M., Wang, Y., Wang, F., Ding, Y., & Song, Y. (2012). The N2pc is increased by perceptual learning but is unnecessary for the transfer of learning. *PLoS ONE*, *7*(4), e34826. https://doi.org/10.1371/journal.pone.0034826

Baluch, F., & Itti, L. (2010). Training top-down attention improves performance on a triple-conjunction search task. *PLoS ONE*, *5*(2), e9127. https://doi.org/10.1371/journal.pone.0009127

Bergen, J. R., & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man and Cybernetics*, *SMC-13*(5), 857–863. https://doi.org/10.1109/TSMC.1983.6313080

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Bueichekú, E., Miró-Padilla, A., & Ávila, C. (2019). Resting-state fMRI detects the effects of learning in short term: A visual search training study. *Human Brain Mapping*, *40*(9), 2787–2799. https://doi.org/10.1002/hbm.24560

Bueichekú, E., Miró-Padilla, A., Palomar-García, M.-Á., Ventura-Campos, N., Parcet, M.-A., Barrós-Loscertales, A., & Ávila, C. (2016). Reduced posterior parietal cortex activation after training on a visual search task. *NeuroImage*, *135*, 204–213. https://doi.org/10.1016/j.neuroimage.2016.04.059

Clark, K., Appelbaum, L. G., Berg, B. van den, Mitroff, S. R., & Woldorff, M. G. (2015). Improvement in visual search with practice: mapping learning-related changes in neurocognitive stages of processing. *Journal of Neuroscience*, *35*(13), 5351–5359. https://doi.org/10.1523/JNEUROSCI.1152-14.2015

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trail EEG dynamics including independent component anlaysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Co-registration of eye movements and EEG in natural reading: analyses and review. *Journal of Experimental Psychology: General*, *140*(4), 552–572. https://doi.org/10.1037/a0023885

Duncan, J. (1989). Boundary conditions on parallel processing in human vision. *Perception*, *18*(4), 457–469. https://doi.org/10.1068/p180457

Duncan, J., & Humphrey, G. W. (1989). Visual serach and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037//0033-295x.96.3.433

Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology*, *3*(4), 674–698. https://doi.org/10.1016/0010-0285(72)90026-6

Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526–535. https://doi.org/10.1016/j.tics.2014.05.005

Ellison, A., & Walsh, V. (1998). Perceptual learning in visual search: some evidence of specificities. *Vision Research*, *38*(3), 333–345. https://doi.org/S0042-6989(97)00195-8

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*,

*9*(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience*, *18*(4), 604–613. https://doi.org/10.1162/jocn.2006.18.4.604

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787–795. https://doi.org/10.3758/BF03206794

Lee, J., Leonard, C. J., Luck, S. J., & Geng, J. J. (2018). Dynamics of feature-based attentional selection during color–shape conjunction search. *Journal of Cognitive Neuroscience*, *30*(12), 1773–1787. https://doi.org/10.1162/jocn

Leonards, U., Rettenbach, R., Nase, G., & Sireteanu, R. (2002). Perceptual learning of highly demanding visual search tasks. *Vision Research*, *42*(18), 2193–2204. https://doi.org/10.1016/s0042-6989(02)00134-7

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*(5), 635–654. https://doi.org/10.1016/j.visres.2007.12.009

Li, H.-H., Pan, J., & Carrasco, M. (2021). Different computations underlie overt presaccadic and covert spatial attention. *Nature Human Behaviour*, 1–14. https://doi.org/10.1038/s41562-021-01099-4

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.x

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*(11), 8. https://doi.org/10.1167/9.11.8

Moore, T., & Fallah, M. (2001). Control of eye movements and spatial attention. *Proceedings of the National Academy of Sciences*, *98*(3), 1273–1276. https://doi.org/10.1073/pnas.98.3.1273

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897x00366

Plöchl, M., Ossandón, J. P., & König, P. (2012). Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, *6*, 278. https://doi.org/10.3389/fnhum.2012.00278

Posner, M. I. (1980). Orienting or attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Qu, Z., Hillyard, S. A., & Ding, Y. (2017). Perceptual learning induces persistent attentional capture by nonsalient shapes. *Cerebral Cortex*, *27*(2), 1512–1523. https://doi.org/10.1093/cercor/bhv342

Reavis, E. A., Frank, S. M., Greenlee, M. W., & Tse, P. U. (2016). Neural correlates of context-dependent feature conjunction learning in visual search tasks. *Human Brain Mapping*, *37*(6), 2319–2330. https://doi.org/10.1002/hbm.23176

Reavis, E. A., Frank, S. M., & Tse, P. U. (2018). Learning efficient visual search for stimuli containing diagnostic spatial configurations and color-shape conjunctions. *Attention, Perception, & Psychophysics*, *80*(5), 1110–1126. https://doi.org/10.3758/s13414-018-1516-9

Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, *3*(3), 264–269. https://doi.org/10.1038/72979

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*(1), 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Sirenteanu, R., & Rettenbach, R. (2000). Perceptual learning in visual search generalizes over task, locations, and eyes. *Vision Research*, *40*(21), 2925–2949. https://doi.org/10.1016/S0042-6989(00)00145-0

Sireteanu, R., & Rettenbach, R. (1995). Perceptual learning in visual search: Fast, enduring, but non-specific. *Vision Research*, *35*(14), 2037–2043. https://doi.org/10.1016/0042-6989(94)00295-W

Su, Y., Lai, Y., Huang, W., Tan, W., Qu, Z., & Ding, Y. (2014). Short-term perceptual learning in visual conjunction search. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1415–1424. https://doi.org/10.1037/a0036337

Talcott, T. N., & Gaspelin, N. (2021). Eye movements are not mandatorily preceded by the N2pc component. *Psychophysiology*, *58*(6), e13821. https://doi.org/10.1111/psyp.13821

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 12–31. https://doi.org/10.1037/0096-1523.10.1.12

Weaver, M. D., van Zoest, W., & Hickey, C. (2017). A temporal dependency account of attentional inhibition in oculomotor control. *NeuroImage*, *147*, 880–894. https://doi.org/10.1016/j.neuroimage.2016.11.004

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168. https://doi.org/10.1016/j.tics.2011.02.005

Table 1. Statistics of Experiment 1. Repeated-measures ANOVAs with session (pretest and posttest) and target type (C+O+, C+O– and C–O+) as factors were performed.

A. Statistics on behavioral d' and RTs.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| d' | F(1,14)=28.54 | <0.001 | 0.67 | F(2,28)=37.07 | <0.001 | 0.73 | F(1.45,20.34)=1.70 | 0.20 | 0.11 |
| RTs | F(1,14)=82.29 | <0.001 | 0.85 | F(2,28)=4.15 | <0.05 | 0.23 | F(2,28)=2.77 | 0.08 | 0.16 |

B. Statistics on eye movement measurements.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| **Zero-saccade trials** | | | | | | | | | |
| Proportion of trials | F(1,14)=5.42 | <0.05 | 0.28 | F(1.09,15.20)=3.11 | 0.10 | 0.18 | F(1.17,16.32)=3.46 | 0.08 | 0.20 |
| **Single on-target-saccade trials (target-present only)** | | | | | | | | | |
| Proportion of trials | F(1,14)=45.67 | <0.001 | 0.77 | F(2,28)=2.29 | 0.12 | 0.14 | F(2,28)=1.23 | 0.31 | 0.08 |
| **Other trials** | | | | | | | | | |
| Search initiation time | F(1,14)=5.74 | <0.05 | 0.29 | F(2,28)=1.81 | 0.18 | 0.11 | F(2,28)=0.19 | 0.83 | 0.01 |
| Number of fixations | F(1,14)=106.63 | <0.001 | 0.88 | F(2,28)=8.18 | <0.01 | 0.37 | F(2,28)=1.28 | 0.29 | 0.08 |
| Mean fixation duration | F(1,14)=0.07 | 0.80 | 0.01 | F(2,28)=0.59 | 0.56 | 0.04 | F(2,28)=0.20 | 0.82 | 0.01 |
| Scanning time | F(1,14)=96.69 | <0.001 | 0.87 | F(2,28)=5.36 | <0.05 | 0.28 | F(2,28)=2.20 | 0.13 | 0.14 |
| Verification time | F(1,11)=3.28 | 0.10 | 0.23 | F(2,22)=1.15 | 0.33 | 0.09 | F(2,22)=2.22 | 0.13 | 0.17 |

Table 2. Statistics of Experiment 2. Repeated-measures ANOVAs with session (pretest and posttest) and target type (C+O+, C+O– and C–O+) as factors were performed.

A. Statistics on behavioral d' and RTs.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| d' | F(1,14)=42.73 | <0.001 | 0.75 | F(2,28)=6.72 | <0.01 | 0.32 | F(1.29,18.12)=9.40 | <0.01 | 0.40 |
| RTs | F(1,14)=50.42 | <0.001 | 0.78 | F(1.21,16.90)=4.35 | <0.05 | 0.24 | F(1.24,17.42)=9.03 | <0.01 | 0.39 |

B. Statistics on eye movement measurements.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| **Zero-saccade trials** | | | | | | | | | |
| Proportion of trials | F(1,14)=0.85 | 0.37 | 0.06 | F(1.27,17.82)=0.83 | 0.40 | 0.06 | F(1.36,19.07)=1.73 | 0.21 | 0.11 |
| **Single on-target-saccade trials (target-present only)** | | | | | | | | | |
| Proportion of trials | F(1,14)=8.60 | <0.05 | 0.38 | F(2,28)=1.82 | 0.18 | 0.12 | F(1.39,19.42)=2.03 | 0.17 | 0.13 |
| **Other trials** | | | | | | | | | |
| Search initiation time | F(1,14)=0.41 | 0.53 | 0.03 | F(2,28)=1.35 | 0.28 | 0.09 | F(2,28)=0.09 | 0.92 | 0.01 |
| Number of fixations | F(1,14)=90.77 | <0.001 | 0.87 | F(1.28,17.92)=6.89 | <0.05 | 0.33 | F(2,28)=7.32 | <0.01 | 0.34 |
| Mean fixation duration | F(1,14)=0.02 | 0.90 | 0.00 | F(2,28)=2.28 | 0.12 | 0.14 | F(2,28)=4.57 | <0.05 | 0.25 |
| Scanning time | F(1,14)=53.31 | <0.001 | 0.79 | F(1.21,16.97)=5.49 | <0.05 | 0.28 | F(1.44,20.12)=10.44 | <0.01 | 0.43 |
| Verification time | F(1,14)=0.91 | 0.36 | 0.06 | F(2,28)=1.75 | 0.19 | 0.11 | F(2,28)=1.76 | 0.19 | 0.11 |

Table 3. Statistics of Experiment 3. Repeated-measures ANOVAs with session (pretest and posttest) and target type (C+O+, C+O– and C–O+) as factors were performed.

A. Statistics on behavioral d' and RTs.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| d' | F(1,14)=17.20 | <0.001 | 0.55 | F(2,28)=0.38 | 0.69 | 0.03 | F(2,28)=11.15 | <0.001 | 0.44 |
| RTs | F(1,14)=62.40 | <0.001 | 0.82 | F(2,28)=5.15 | <0.05 | 0.27 | F(1.41,19.70)=3.20 | 0.08 | 0.19 |

B. Statistics on eye movement measurements.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| **Zero-saccade trials** | | | | | | | | | |
| Proportion of trials | F(1,14)=3.90 | 0.07 | 0.22 | F(1.37,19.19)=0.61 | 0.50 | 0.04 | F(2,28)=1.17 | 0.32 | 0.08 |
| **Single on-target-saccade trials (target-present only)** | | | | | | | | | |
| Proportion of trials | F(1,14)=40.16 | <0.001 | 0.74 | F(2,28)=1.08 | 0.35 | 0.07 | F(2,28)=0.29 | 0.75 | 0.02 |
| **Other trials** | | | | | | | | | |
| Search initiation time | F(1,14)=0.46 | 0.51 | 0.03 | F(2,28)=0.81 | 0.45 | 0.05 | F(2,28)=2.70 | 0.08 | 0.16 |
| Number of fixations | F(1,14)=48.58 | <0.001 | 0.78 | F(2,28)=10.62 | <0.001 | 0.43 | F(2,28)=3.09 | 0.06 | 0.18 |
| Mean fixation duration | F(1,14)=6.33 | <0.05 | 0.31 | F(1.43,20.07)=0.33 | 0.65 | 0.02 | F(2,28)=0.13 | 0.88 | 0.01 |
| Scanning time | F(1,14)=43.83 | <0.001 | 0.76 | F(2,28)=8.22 | <0.01 | 0.37 | F(1.42,19.92)=2.79 | 0.10 | 0.17 |
| Verification time | F(1,14)=0.74 | 0.40 | 0.05 | F(1.34,18.75)=0.84 | 0.40 | 0.06 | F(2,28)=2.70 | 0.08 | 0.16 |

Table 4. Statistics of Experiment 4. Repeated-measures ANOVAs with session (pretest and posttest) and target type (C+O+, C+O– and C–O+) as factors were performed.

A. Statistics on behavioral d' and RTs.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| d' | F(1,14)=22.07 | <0.001 | 0.61 | F(2,28)=1.13 | 0.28 | 0.09 | F(2,28)=1.23 | 0.31 | 0.08 |
| RTs | F(1,14)=16.55 | <0.01 | 0.54 | F(2,28)=1.23 | 0.31 | 0.08 | F(2,28)=0.21 | 0.81 | 0.01 |

B. Statistics on eye movement measurements.

| | **Session** | | | **Target type** | | | **Interaction** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ | *F* | *p* | $\eta_p^2$ |
| **Zero-saccade trials** | | | | | | | | | |
| Proportion of trials | F(1,14)=0.58 | 0.46 | 0.04 | F(1.08,15.17)=1.74 | 0.21 | 0.11 | F(2,28)=0.66 | 0.52 | 0.05 |
| **Single on-target-saccade trials (target-present only)** | | | | | | | | | |
| Proportion of trials | F(1,14)=1.40 | 0.26 | 0.09 | F(2,28)=0.22 | 0.80 | 0.02 | F(2,28)=1.57 | 0.23 | 0.10 |
| **Other trials** | | | | | | | | | |
| Search initiation time | F(1,14)=0.18 | 0.68 | 0.01 | F(2,28)=3.07 | 0.06 | 0.18 | F(2,28)=0.67 | 0.52 | 0.05 |
| Number of fixations | F(1,14)=14.70 | <0.01 | 0.51 | F(2,28)=5.22 | <0.05 | 0.27 | F(2,28)=0.21 | 0.81 | 0.01 |
| Mean fixation duration | F(1,14)=0.25 | 0.63 | 0.02 | F(2,28)=3.85 | <0.05 | 0.22 | F(2,28)=0.29 | 0.75 | 0.02 |
| Scanning time | F(1,14)=16.05 | <0.01 | 0.53 | F(2,28)=2.19 | 0.13 | 0.14 | F(2,28)=0.07 | 0.94 | 0.00 |
| Verification time | F(1,14)=0.41 | 0.53 | 0.03 | F(2,28)=0.17 | 0.85 | 0.01 | F(2,28)=3.43 | <0.05 | 0.20 |

Table 5. Statistics of Experiment 5. Paired t-tests (trained target vs. untrained target) were performed.

A. Statistics on behavioral d' and RTs.

| | **Target type** | | |
|---|---|---|---|
| | *t* | *p* | *Cohen's d* |
| d' | t(19)=5.38 | <0.001 | 1.20 |
| RTs | t(19)=-9.17 | <0.001 | 2.05 |

B. Statistics on eye movement measurements.

| | **Target type** | | |
|---|---|---|---|
| | *t* | *p* | *Cohen's d* |
| **Zero-saccade trials** | | | |
| Proportion of trials | t(19)=3.38 | <0.01 | 0.76 |
| **Single on-target-saccade trials (target-present only)** | | | |
| Proportion of trials | t(19)=3.51 | <0.01 | 0.78 |
| **Other trials** | | | |
| Search initiation time | t(19)=2.42 | <0.05 | 0.54 |
| Number of fixations | t(19)=-9.47 | <0.001 | 2.12 |
| Mean fixation duration | t(19)=-1.36 | 0.19 | 0.30 |
| Scanning time | t(19)=-9.31 | <0.001 | 2.08 |
| Verification time | t(19)=--2.04 | 0.06 | 0.46 |

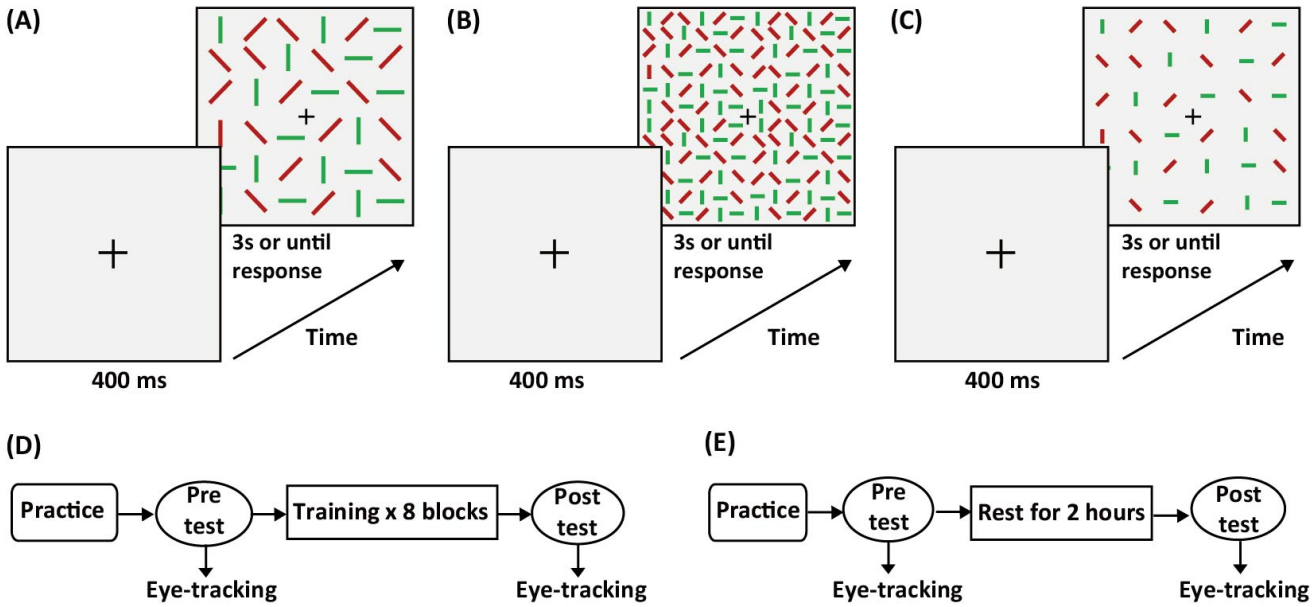


Figure 1. Stimuli and typical trials of the conjunction visual search task in (A) Experiment 1, (B) Experiments 2 and 4, and (C) Experiment 3. (D) Procedure in Experiments 1, 2, and 3. (E) Procedure in Experiment 4.

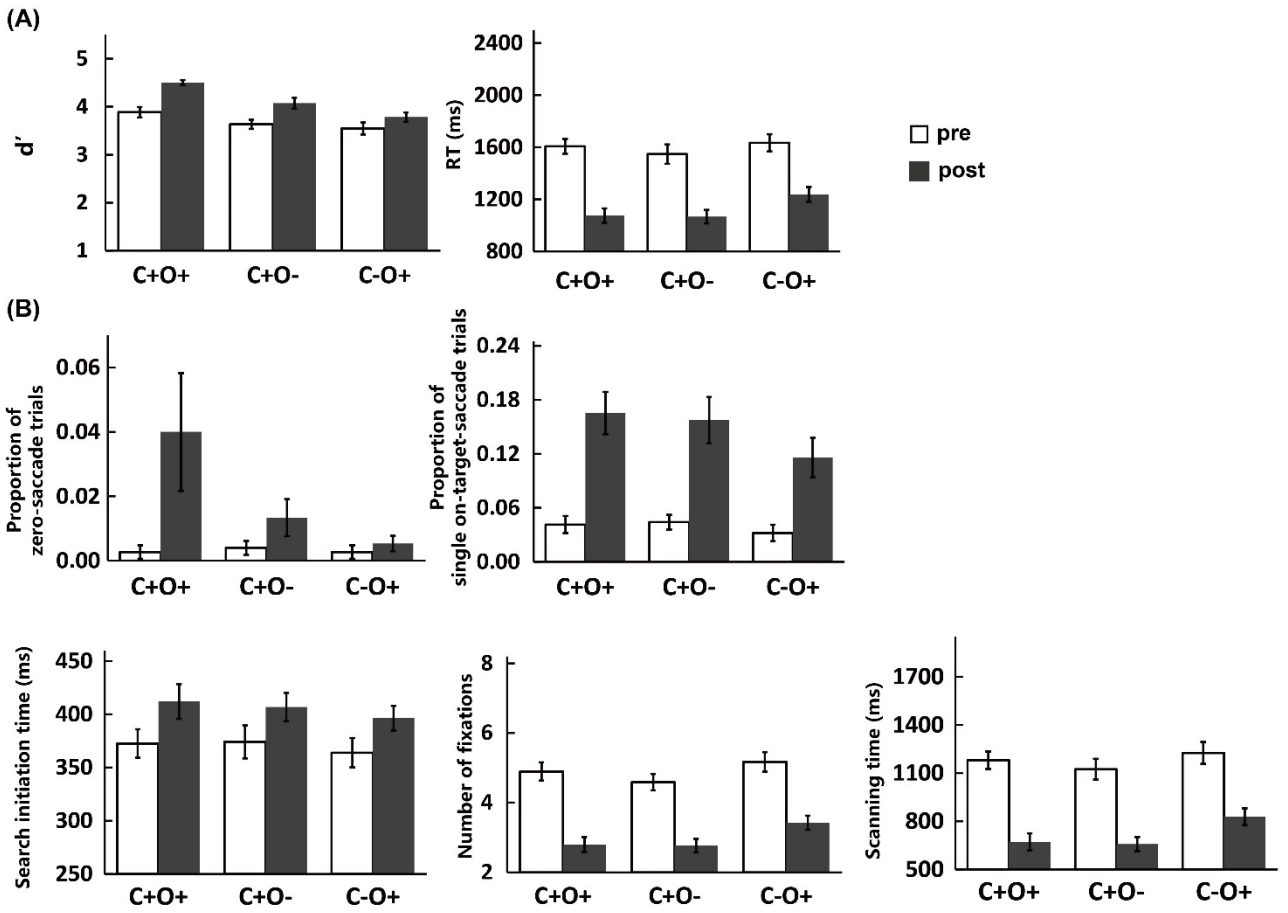


Figure 2. Results of Experiment 1. (A) Behavioral results for d' and RTs. See Table 1A for statistical results. (B) Eye movement results for the proportion of zero-saccade trials, proportion of single on-target-saccade trials, and other trials' search initiation time, number of fixations, and scanning time. See Table 1B for statistical results. Error bars represent standard errors of the mean across participants.

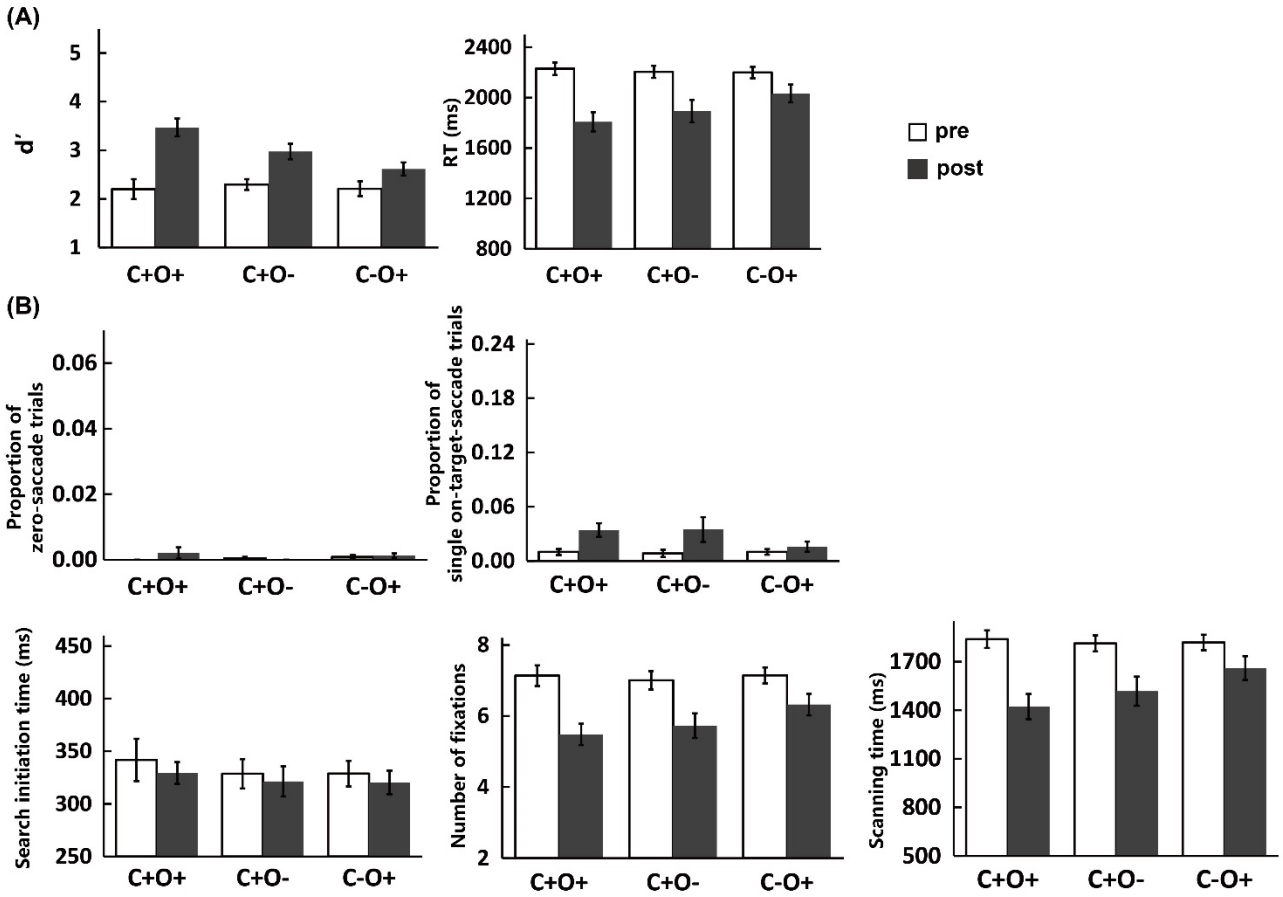


Figure 3. Results of Experiment 2. (A) Behavioral results for d' and RTs. See Table 2A for statistical results. (B) Eye movement results for the proportion of zero-saccade trials, proportion of single on-target-saccade trials, and other trials' search initiation time, number of fixations, and scanning time. See Table 2B for statistical results. Error bars represent standard errors of the mean across participants.

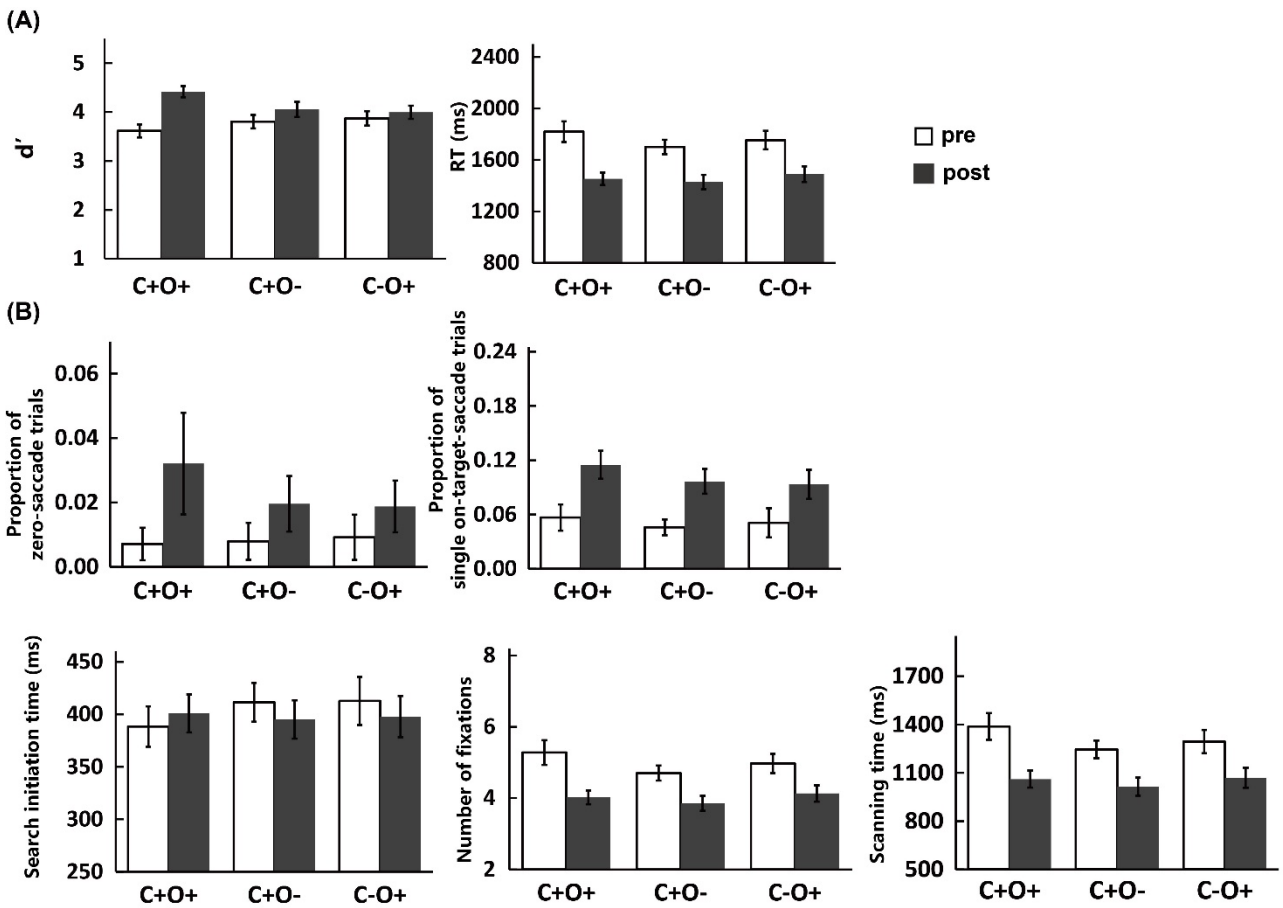


Figure 4. Results of Experiment 3. (A) Behavioral results for d' and RTs. See Table 3A for statistical results. (B) Eye movement results for the proportion of zero-saccade trials, proportion of single on-target-saccade trials, and other trials' search initiation time, number of fixations, and scanning time. See Table 3B for statistical results. Error bars represent standard errors of the mean across participants.

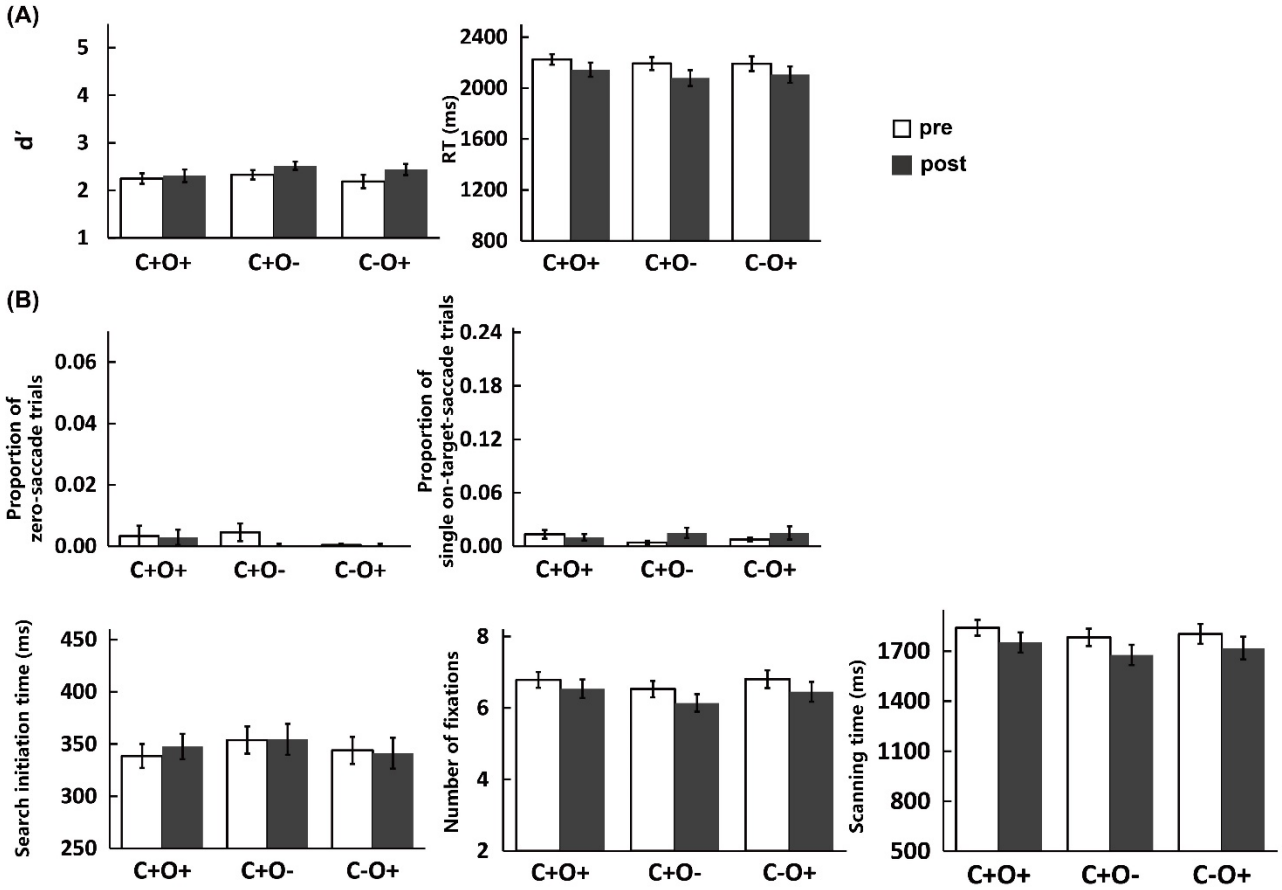


Figure 5. Results of Experiment 4. (A) Behavioral results for d' and RTs. See Table 4A for statistical results. (B) Eye movement results for the proportion of zero-saccade trials, proportion of single on-target-saccade trials, and other trials' search initiation time, number of fixations, and scanning time. See Table 4B for statistical results. Error bars represent standard errors of the mean across participants.

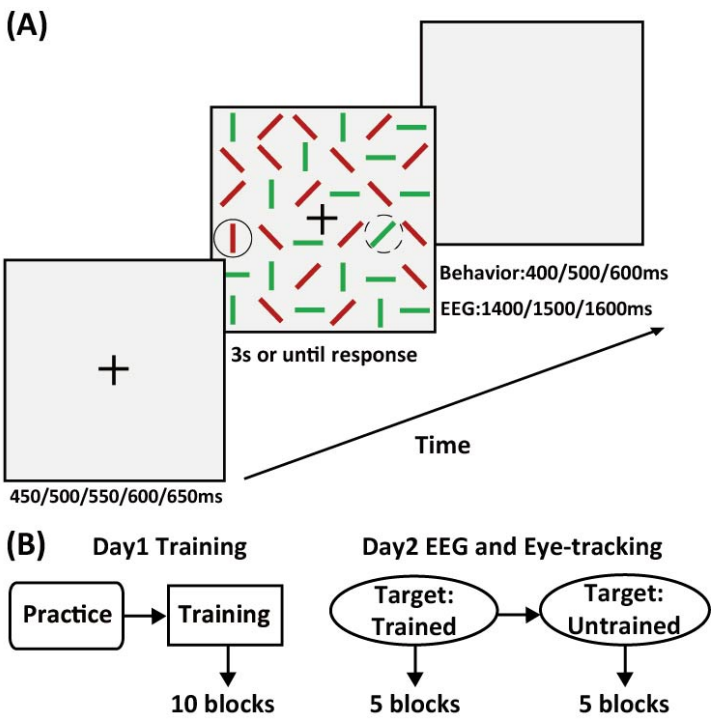


Figure 6. Experimental design of Experiment 5. (A) Stimulus and trial sequence. Solid line circle (trained target) and dashed line circles (untrained target served as distractor) are drawn for illustration purpose of training session and were not shown during the experiment. (B) Experimental procedure.

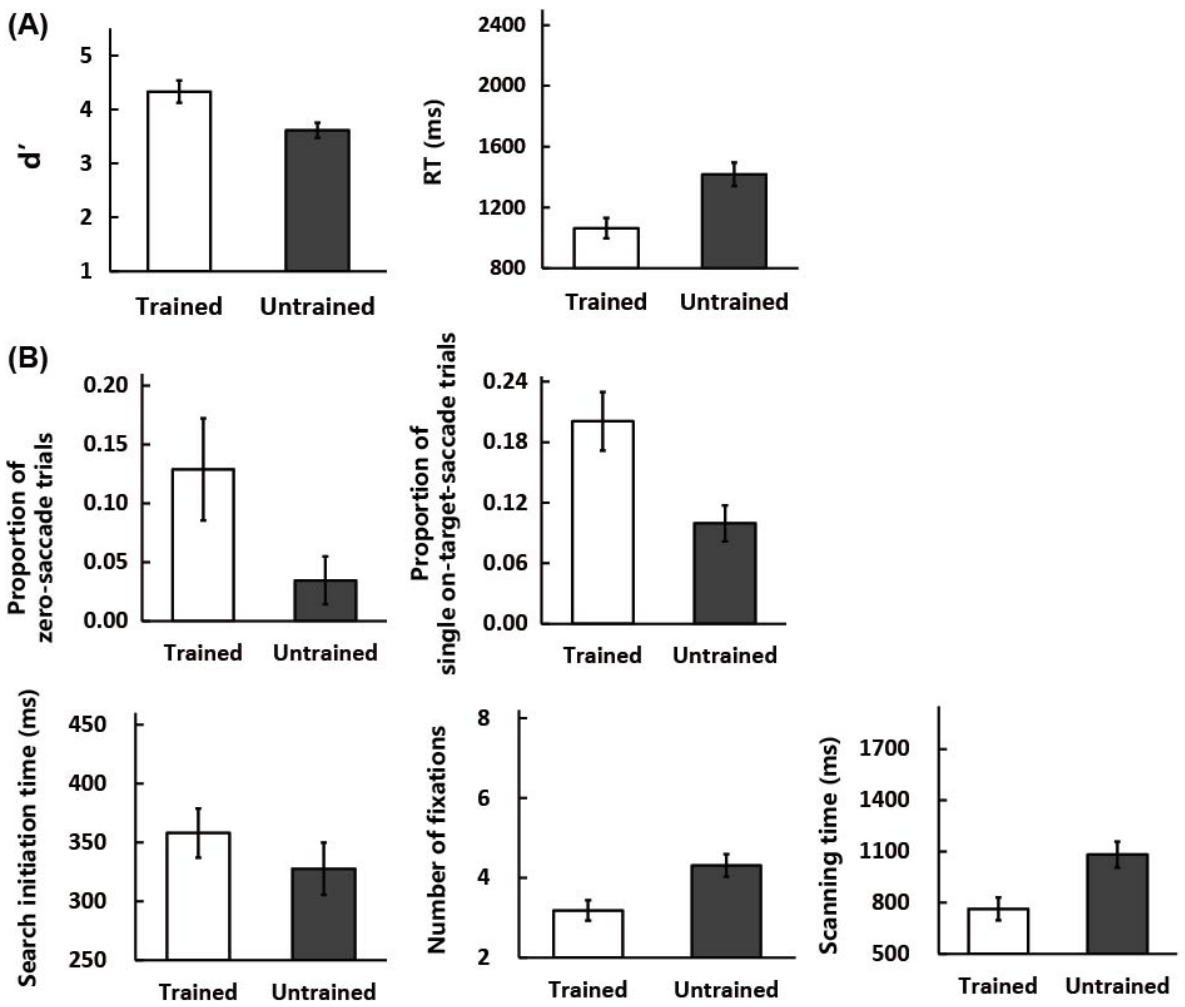


Figure 7. Behavioral and eye movement results of Experiment 5 for trials with trained and untrained targets. (A) Behavioral results for d' and RTs. See Table 5A for statistical results. (B) Eye movement results for the proportion of zero-saccade trials, proportion of single on-target-saccade trials, and other trials' search initiation time, number of fixations, and scanning time. See Table 5B for statistical results. Error bars represent standard errors of the mean across participants.

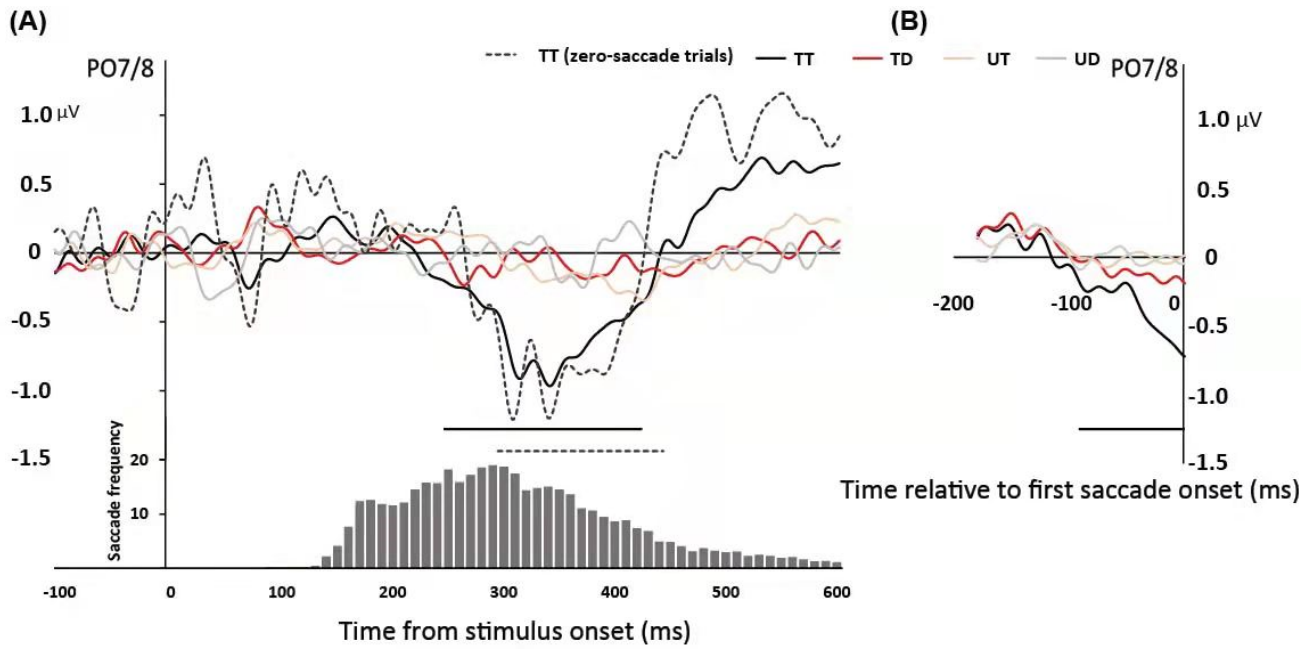


Figure 8. EEG results of Experiment 5. (A) Stimulus-locked N2pc components are shown for conditions in which the trained stimulus and untrained stimulus served as target or distractor. Black dash curve is the N2pc for TT condition of the zero-saccade trials. Frequency distribution of first saccade onset time aggregated across conditions and participants are presented below the ERP curves. (B) Saccade-locked N2pc components are shown for four conditions with eye movements. Periods of significant N2pc amplitudes were shown as straight lines below the ERPs. TT: trained stimulus served as target; TD: trained stimulus served as distractor; UT: untrained stimulus served as target; UD: untrained stimulus served as distractor.