# Inhibition of return as a foraging facilitator in visual search: Evidence from long-term training

Ai-Su Li[1,2†], Yan Li[3†], Xun He[4*], Yang Zhang[1*]

[1]Department of Psychology, Soochow University, Suzhou, China

[2]Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[3]Department of Psychology, Tsinghua University, Beijing, China

[4]Department of Psychology, Bournemouth University, Poole, UK

Running Head: Search efficiency and IOR during training

[†]Ai-Su Li and Yan Li contributed equally to this study and are co-first authors.

[*]Corresponding authors:

Yang Zhang (Department of Psychology, Soochow University, WenJing Street 1, 215000 Suzhou, China; e-mail: yzhangpsy@suda.edu.cn; tel: +86 181 36777977)

Xun He (Department of Psychology, Bournemouth University, Fern Barrow, Poole, Dorset BH12 5BB, UK; e-mail: xhe@bournemouth.ac.uk; tel: +44 1202 965148).

**Abstract**

Inhibition of return (IOR) discourages visual attention from returning to previously attended locations, and has been theorized as a mechanism to facilitate foraging in visual search by inhibitory tagging of inspected items. Previous studies using visual search and probe detection tasks (i.e., the probe-following-search paradigm) found longer RTs for probes appearing at the searched locations than probes appearing at novel locations. This IOR effect was stronger in serial than parallel search, favoring the foraging facilitator hypothesis. However, evidence for this hypothesis was still lacking because no attempt was made to study how IOR would change when search efficiency gradually improves. The current study employed the probe-following-search paradigm and long-term training to examine how IOR varied following search efficiency improvements across training days. According to the foraging facilitator hypothesis, inhibitory tagging is an aftereffect of attentional engagement. Therefore, when attentional engagement in a visual search task is reduced via long-term training, the strength of inhibitory tagging would decrease and thus predicts a reduced IOR effect. Consistent with this prediction, two experiments consistently showed that IOR decreased while search efficiency improved through training, although IOR reached the floor more quickly than search efficiency. These findings support the notion that IOR facilitates search performance via stronger inhibitory tagging in more difficult visual search.

**Keywords:** inhibition of return; visual search; foraging facilitator; inhibitory tagging; training

**Introduction**

Efficiently searching for a target object in a cluttered environment (e.g., looking for your old friend at a train station) is a fundamental skill of the human visual system (Ma et al., 2011; Najemnik & Geisler, 2005; Wolfe, 2003; Zhang et al., 2021). To achieve optimal search performance, it is necessary not to frequently re-visit the locations that have been previously (Gilchrist & Harvey, 2000; Najemnik & Geisler, 2005). Posner and Cohen (1984) demonstrated that shortly after attention has moved away from a location, manual responses and eye movements to stimuli appearing at the previously attended loci were delayed compared with those at new locations. They proposed that such inhibition effect encourages the orienting of attention towards novel items to maximize sampling of the visual environment. Posner et al. (1985) later named this inhibition effect *inhibition of return* (IOR).

Klein (1988) proposed the *foraging facilitator hypothesis*, which suggests that IOR facilitates performance in visual search tasks that require the serial deployment of attention by attaching inhibitory tags to inspected spatial locations and discouraging attention from being oriented back towards the already tagged locations if the search needs to continue. To test this hypothesis, Klein (1988) developed the probe-following-search paradigm, which includes a visual search task (to generate inhibitory tags) and the following probe detection task (to measure inhibitory tags) in each trial. In visual search involving serial attentional deployment, if inhibitory tagging takes place at the inspected locations in the search display, detection of the probe stimulus appearing at any of these locations should be slower due to attention being discouraged from returning to these locations. Therefore, the hypothesis suggests longer reaction times (RTs) to the probes presented at the tagged distractor locations

(i.e., the *on-probe* condition) than to the probes appeared at blank locations (i.e., the *off-probe* condition). The RT difference between these two conditions, known as the *on-probe cost*, provides a measure of IOR in visual search tasks. Furthermore, to control for potential influence from expectancy or forward masking on probe detection, an easy parallel search task (followed by the same probe detection task) was employed as a control condition. In parallel search (pop-out search), the target is salient and captures attention, and is detected without sequential inspection or attentional shifts, thus not involving any inhibitory tagging. Therefore, the on-probe cost in parallel search would be smaller than that in serial search. Klein (1988) showed the predicted pattern and supported the foraging facilitator hypothesis.

However, the foraging facilitator hypothesis was challenged by two studies (Klein & Taylor, 1994; Wolfe & Pokorny, 1990), which reported failures to replicate the findings of Klein (1988). After the search task was completed in these studies, all search items were removed from the display before the detection probe appeared. Despite the finding of longer RTs to on-probes than to off-probes, there was no difference in the on-probe cost between parallel search and serial search, showing a forward masking effect instead of inhibitory tagging. This dispute remained unresolved until researchers found that an enlarged IOR in serial search could only be observed when the search array was visible when the probe appeared to enable marking previously attended locations (Müller & von Mühlenen, 2000; Takeda & Yagi, 2000). This conclusion is consistent with the foraging facilitator hypothesis because landmarks usually remain in a natural foraging scene. The sudden disappearance of some landmarks may indicate highly salient environmental changes requiring attentional focus instead of attentional inhibition.

According to the foraging facilitator hypothesis, IOR should be weaker in efficient search tasks than in inefficient search tasks (Klein, 1988; Wang & Klein, 2010). However, little research was conducted to investigate how IOR changes when search efficiency varies. One way to manipulate search efficiency is to employ serial search tasks of various search difficulties. However, this approach requires using different stimuli across tasks to attain the various expected search efficiencies, rendering the results confounded by differences in the physical stimuli. Therefore, a better approach is to adopt a training procedure capable of gradually increasing the search efficiency while keeping the stimuli identical in the process. It is also important to note that, in Klein's (1988) study, the stimuli differed between the serial search task and the parallel search task. Therefore, it is desirable to keep the physical stimuli as similar as possible between the serial and parallel search tasks in the training procedure to achieve a better-matched control condition.

The current study combined the probe-following-search paradigm (Klein, 1988) and a training procedure. IOR was quantified as the difference between the on-probe cost in serial search and that in parallel search. In particular, the focus was on the IOR in the target-absent condition because it is a more reliable measure. Inefficient (serial) search is executed with a serial and self-terminating strategy, checking search items one by one until the target is detected (target-present condition) or when a quitting signal reaches a threshold after all the search items with high probabilities of being the search target are exhausted (target-absent condition) (Chun & Wolfe, 1996; Wolfe, 2021). In the target-present trials, the search, on average, will be terminated after half of the search items have been checked, meaning that there are great chances that the probe is presented at a location previously occupied by a

search item yet never searched, leading to an unreliable estimation of IOR. In contrast, IOR measurement in the target-absent trials is more reliable because all probable search items will be attended to in the visual search. The search-terminating threshold is adaptive according to the search requirements and the observer's tolerance to search errors (e.g., Wolfe, 2021). Even though the search may not exhaust all items when the search target is absent, most items will be checked when a high search accuracy is reached. Therefore, in the current study, the on-probe cost difference between the serial and parallel search tasks in the target-absent trials was taken as the IOR measure.

According to the foraging facilitator hypothesis, inhibitory tagging is applied to the inspected locations to discourage attention from returning to these locations (until the search target is found), producing the IOR effect and subsequently facilitating the search performance. When the visual search gradually becomes more efficient (easier) in a training procedure, visual attention may remain qualitatively serial, but less attentional engagement is involved over time. This will lead to weaker inhibitory tagging, in turn, reduce the IOR effect. Therefore, the foraging facilitator hypothesis (inhibitory tagging) will be supported if the data shows a decrease in IOR over training days as the search efficiency increases.

It is important to note that the above prediction is based on the specific manipulation employed. In the current training protocol, the level of inhibitory tagging is determined by search difficulty (measured by search efficiency), so that more efficient search is predicted to generate weaker inhibitory tagging (thus weaker IOR), showing a negative relationship between search efficiency and IOR. In the extreme case of pop-out search (super-efficient), no sequential attentional shift is involved, thus the IOR effect should be absent (Klein, 1988).

However, an alternative prediction suggests that, given that IOR is a foraging facilitator, an increased IOR will facilitate search performance and as a result increase search efficiency. Opposite to the current hypothesis, this would predict a positive relationship between search efficiency and IOR. Note that this prediction requires directly manipulating the strength of IOR. Currently, it is unclear how to manipulate IOR (either via volitional control or via other variables), not to mention that the current study did not manipulate IOR directly, rendering this alternative prediction inapplicable to the current study. Therefore, it is hypothesized that the IOR and search efficiency would show a negative relationship in the current study.

## Experiment 1

### Method

*Participants*

Twenty naïve volunteers (18 females and 2 males, $M_{age}$ = 20.50 years, $SE_{age}$ = 0.51, all right-handed) took part in Experiment 1 after obtaining informed consent. Four participants were excluded from the analysis due to the absence of IOR in the target-absent condition on the first training day. All participants had normal or corrected-to-normal vision and had no experience with similar tasks before. All procedures were approved by the Academic Committee of Department of Psychology at Soochow University.

The sample size was determined with a power analysis using MorePower (Campbell & Thompson, 2012), with $\alpha$ = .05, $\beta$ = .2, and an effect size of $\eta^2$ = 0.14 for the interaction effect in a two-way (search target presence × training days) repeated-measures analysis of variance (RM-ANOVA) on IOR. The result suggested a sample size of 16.

*Apparatus*

The participants were tested individually in a dimly lit and sound-attenuated room. All stimuli were presented against a gray background on an 18-inch CRT monitor (ViewSonic P220f) with a resolution of 1,024×768 and a refresh rate of 100 Hz (Zhang et al., 2018). The participants sat at a viewing distance of 80 cm with their chins resting on a chin rest. Stimulus presentation and response recording were controlled by Psychtoolbox-3 (Kleiner et al., 2007) in MATLAB 2011a, running on a Xeon E5-2630 PC equipped with an NVIDIA GTX750 graphics card. Responses were collected with a standard computer keyboard.

*Design*

The current study used a within-participants design with four factors, namely training days (days 1 to 7), search tasks (serial vs. parallel), search target presence (present vs. absent), and probe types (on-probe, off-probe, and no probe). A probe was presented at a location previously occupied by a search item in the on-probe condition and at a previously empty location in the off-probe condition. The probe was absent in the no probe condition.

*Procedure*

The participants were trained for seven consecutive days with the probe-following-search task. Each participant completed two visual search sessions on each training day, one parallel search and the other serial search. The completion order of the sessions was counterbalanced across participants but remained identical across training days for each participant. Each

session consisted of six 100-trial blocks separated by short breaks, with the two sessions being separated by a 10-minute break. At the beginning of each session, participants performed 10 to 20 practice trials. In sum, there were 1,200 trials in total on each training day. In each session (parallel or serial), there were 300 trials in each target-presence condition (target-present or target-absent). In the following probe task, out of these 300 trials, 120 were on-probe trials, 120 were off-probe trials, and 60 were no-probe trials. These large trial numbers were adopted following previous studies (Li et al., 2018; Li et al., 2017; Lin et al., 2020) because it was proposed that a trial count of 100 or more per condition is desirable for establishing reliable mean values (Vanselst & Jolicoeur, 1994).

Figure 1 illustrates the trial sequence and stimuli. Each trial began with a black fixation cross (0.4° × 0.4°, 0.08° line thickness), which was presented centrally and remained visible throughout the trial. After 1,500 ms, an array of 8 search items appeared in two random quadrants out of the four quadrants of the display, with four items in each displayed quadrant. This display layout was adopted to produce more substantial IOR magnitudes. It has been found with a cueing paradigm that, when visual targets appeared after a long delay (800 ms) from cue onsets, a gradient of RTs was found throughout the visual field, showing a gradual change from IOR in the cued quadrant to facilitation in the diagonal quadrant (Bennett & Pratt, 2001). Each item was a black outlined square (0.45° × 0.45°, 0.08° line thickness) with a gap (0.12°) on one side. Each quadrant extended 6.1° horizontally and vertically from the fixation cross. Within each quadrant containing search items, the items were placed randomly in an imaginary grid of 3 × 3, excluding the location next to the fixation cross, with a further ±0.525° random jitter for presentation locations within the grid. This presentation protocol

resulted in a minimum distance of 1.05° between search items and a minimum distance of

2.65° between the items and the fixation (both measured center-to-center).
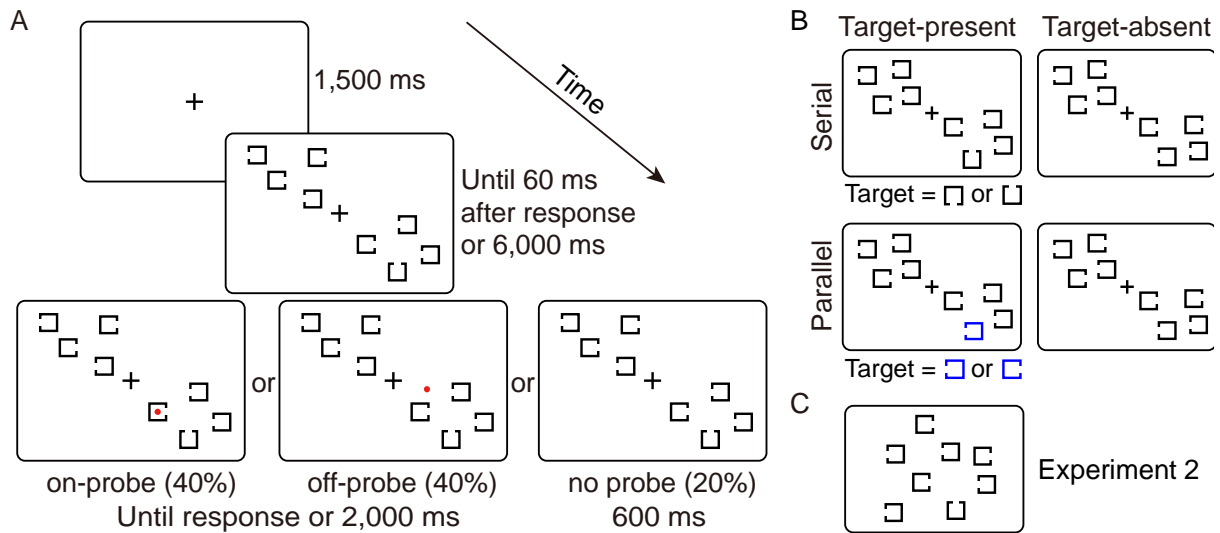


Figure 1. Stimuli in the current study. A) Schematic diagram of stimuli and procedure used in Experiment

1. A visual search task was immediately followed by a probe detection task in each trial. The dot-probe

could be on a search item (on-probe), away from the search items (off-probe), or absent (no probe). B)

Stimuli used in the target-present and target-absent conditions in the serial and parallel search tasks in

Experiment 1. They were displayed in two randomly chosen quadrants out of four (4 items per quadrant).

C) Search layout in Experiment 2. Note that the search items and tasks were the same as in Experiment 1,

except that all the items were displayed in one cluster covering the central area of the display. In both

experiments, the background was gray (shown as white in this figure for illustration purposes).

In the serial search session, the target was a black square with a gap on the top or the

bottom, whereas the distractors were black squares with a gap on the left or the right. In the

parallel search session, all items were squares showing gaps on the left or the right, with the

targets in blue and the distractors in black. Search targets were present in half of the trials and

absent in the other half. Participants were asked to press one of two appointed keys with the right index and middle fingers quickly and accurately to indicate the presence/absence of targets (key mapping counterbalanced across participants) within a response window of 6,000 ms.

In 80% of the trials, a probe stimulus was superimposed on the search display 60 ms after a manual response was given to the search display or 6,000 ms after the onset of the search display if no search response was detected. The proportion of the trials containing the probe was higher than that in Klein's (1988) study (50%) to obtain more data on the probe detection performance. The probe, a red dot of 0.15° in diameter, appeared either at a location occupied by one of the search items (on-probe; 40% of the trials) or at an empty location in a quadrant containing search items (off-probe; 40% of the trials) with equal probabilities. Participants were required to press the space bar with the left index finger as quickly and accurately as they could once detecting the probe. In the on-probe and off-probe trials, the probe stimulus and the search display remained on the screen for 2,000 ms or until a manual response was made (whichever the sooner). In the remaining 20% of the trials in which the probe was absent (no probe), the search display remained in view for 600 ms.

*Data analysis*

For the search task, trials with incorrect search responses (2.24% of total trials) were excluded from the analysis. This very low error rate was not further analyzed, but suggested that the search-terminating threshold in the target-absent trials allowed most search items to be inspected, supporting the argument that the IOR estimation would be much more accurate

in the target-absent trials than in the target-present trials. Trials with correct search responses but RTs shorter than 100 ms or more than 2.5 standard deviations (SDs) away from mean RTs (in each condition, each training day, and each individual) were also rejected (2.25% of total trials).

For the probe detection task, data from probe-present trials were analyzed except for those without detection responses (0.06% of total trials) and those preceded by incorrect search responses (2.17% of total trials; Müller & von Mühlenen, 2000). Furthermore, trials in which probes appeared at the exact search target locations (2.41% of total trials) were rejected because attention might still be dwelling at the target locations when the probe showed up and would diminish IOR (Wang et al., 2010). Lastly, trials with the detection RTs shorter than 100 ms or more than 2.5 SDs away from mean RTs (in each condition, each training day, and each individual) were excluded as well (2.24% of total trials).

Search efficiency was defined as the number of items scanned per second and quantified by dividing the search display size (8 items) by the search RT difference between the serial and parallel search conditions. IOR was quantified by subtracting the on-probe cost in parallel search from that in serial search (Klein, 1988), i.e., $(\text{RT}_{on\text{-serial}} - \text{RT}_{off\text{-serial}}) - (\text{RT}_{on\text{-parallel}} - \text{RT}_{off\text{-parallel}})$. Search efficiency and IOR effects were then submitted to a two-way (search target presence × training days) RM-ANOVA. Greenhouse-Geisser corrected $p$ values ($p_c$) were used in case of sphericity assumption violations.

The current study employed a training protocol to gradually increase the search efficiency and predicted a decrease of IOR over the training days. Therefore, it was crucial to ensure that the search task was difficult enough to engage the inhibitory tagging mechanism

at the start of the training. As such, the participants who did not show the IOR effect in the target-absent condition on the first training day were excluded from the data analysis. The application of this criterion removed four participants. Nevertheless, the analysis including these four participants showed a similar pattern of results (see Supplementary Materials).

**Results**

*Search performance*

Mean RTs in the visual search task are shown in Figure 2 (Panel A). Figure 3 illustrates the search efficiencies in the target-absent (Panel A) and target-present (Panel B) conditions across training days. The RM-ANOVA on search efficiencies showed a significant main effect of search target presence, $F(1, 15) = 83.44$, $p < .0001$, partial $\eta^2 = 0.85$, with the search efficiency doubled in the target-present trials ($18.55 \pm 1.69$ items/sec for $M \pm SE$) than in the target-absent trials ($9.38 \pm 0.91$ items/sec), consistent with previous findings (e.g., Klein, 1988; Müller & von Mühlenen, 2000; Takeda & Yagi, 2000). The main effect of training days was also significant, $F(6, 90) = 34.77$, $p_c < .0001$, partial $\eta^2 = 0.70$, showing a monotonic increase in search efficiency over training days (from $7.01 \pm 0.56$ items/sec on the first training day to $19.32 \pm 2.16$ items/sec on the last training day), mainly from a significant linear contrast, $F(1, 15) = 38.68$, $p < .0001$, partial $\eta^2 = 0.72$.
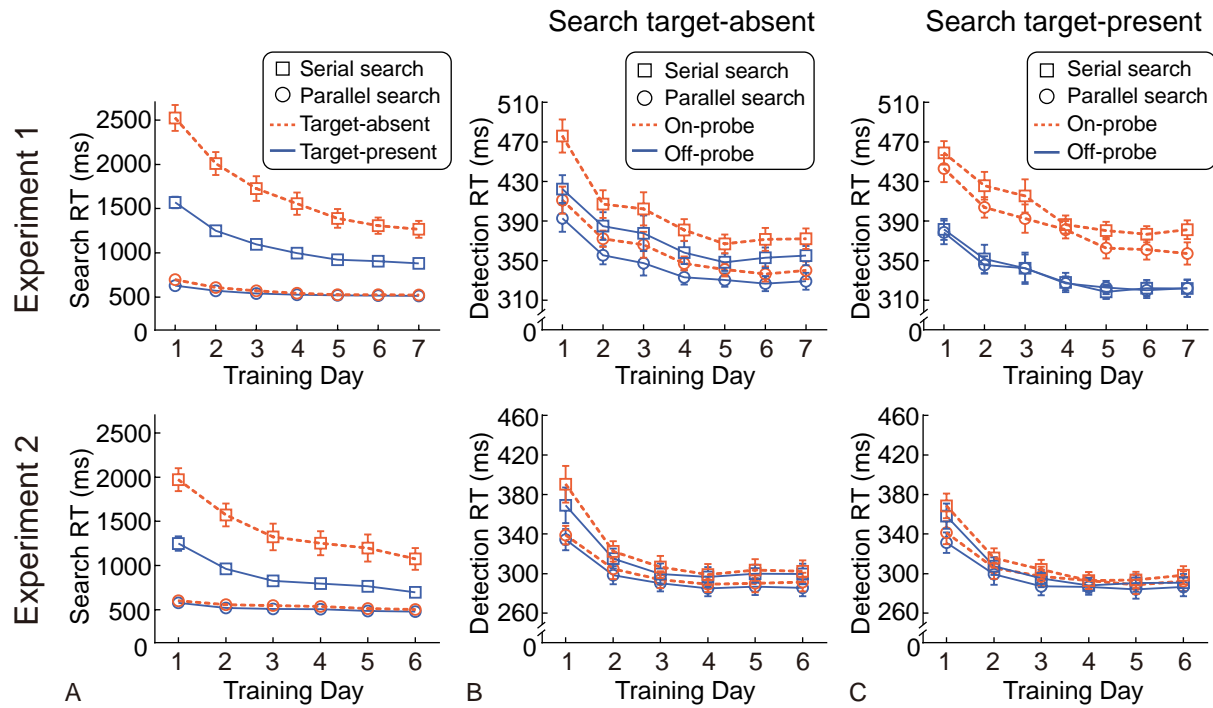
Figure 2. Mean search and probe detection response times (RTs) in Experiments 1 (top row) and 2 (bottom row). Panel A shows the search RTs as a function of training days. Panels B and C respectively show the probe detection RTs in the search target-absent and search target-present conditions against training days. Error bars indicate ±1 SEM.
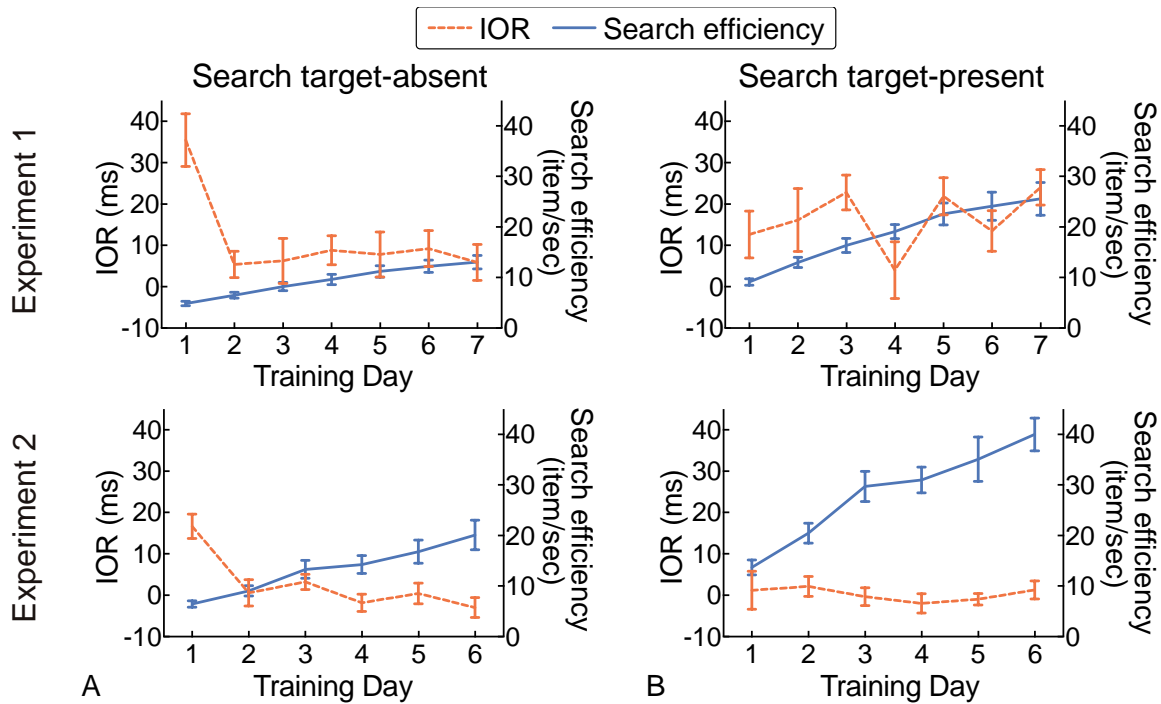
Figure 3. IOR (ms) and search efficiency (items/sec) as a function of training days in the search target-absent (A) and search target-present (B) conditions in Experiments 1 (top row) and 2 (bottom row). Error bars indicate ±1 SEM. Note that the results in the search target-absent condition are more reliable because, in the search target-present condition, there are great chances that the probe was presented at a location that was previously occupied by a search item but was never searched, leading to an unreliable estimation of IOR.

The search target presence × training days interaction was also significant, $F(6, 90) = 9.66$, $p_c = .004$, partial $\eta^2 = 0.39$, suggesting differential search efficiency improvements over time between the target-present and -absent conditions. To statistically assess this, a linear regression was performed on search efficiency over training days for each search target presence condition (present and absent) in each participant. The regression slope would represent the speed of search efficiency improvement over time. A paired-samples $t$-test was then carried out to compare the regression slopes of the target-present condition with those of

the target-absent condition. The result indicated that the slopes were significantly steeper in the target-present condition (2.78 ± 0.52 items/sec/day) than those in the target-absent condition (1.39 ± 0.19 items/sec/day), $t(15) = 3.36$, $p = .004$, Cohen's $d = 0.84$.

*IOR performance*

Mean RTs in the probe detection task are shown in Figure 2 (Panels B and C). The IOR effects in the search target-absent (Panel A) and search target-present (Panel B) conditions are illustrated in Figure 3 as a function of the training days. The IOR performance was analyzed with the same 2 (search target present vs. absent) × 7 (training days) RM-ANOVA. The main effect of training days was significant, $F(6, 90) = 3.16$, $p_c = .02$, partial $\eta^2 = 0.17$. The result did not show a significant main effect of search target presence, $F(1, 15) = 1.13$, $p = .30$, partial $\eta^2 = 0.07$. Furthermore, the interaction reached significance, $F(6, 90) = 5.50$, $p < .0001$, partial $\eta^2 = 0.27$. To further assess this interaction, separate one-way RM-ANOVAs were conducted to investigate the training effects in the target-absent and -present conditions respectively. The results did not reveal a significant difference in IOR across training days in the target-present condition, $F(6, 90) = 2.01$, $p = .07$, partial $\eta^2 = 0.12$. However, a significant training effect was found in the target-absent condition, $F(6, 90) = 8.93$, $p < .0001$, partial $\eta^2 = 0.37$. This result was a combination of a significant linear trend [$F(1, 15) = 12.46$, $p = .003$, partial $\eta^2 = 0.45$] and a significant quadratic trend [$F(1, 15) = 16.58$, $p = .001$, partial $\eta^2 = 0.53$].

The significant quadratic trend suggests the existence of an exponential decline in IOR in search target-absent trials across training days. To further explore this, a linear mixed models

(LMM) analysis (West et al., 2014) was carried out (using the *fitlme* function in MATLAB) on IOR with $e$ to the power of training days arranged in a descending order ($e^7$, $e^6$, $e^5$, …$e^1$) for the fixed effect. By-participants random intercepts and by-participants random slopes for training days were included as random effects. Two models, respectively with and without the fixed effect of training days, were compared (using the *compare* function in MATLAB) to determine the contribution of training days. The result was significant, $\chi^2(1) = 14.77$, $p = .00013$, confirming an exponential decrease of the IOR across training days.

Following the finding of the exponential decrease of the IOR over training, further tests were run to establish the learning curve of the IOR. With seven days in total, running all pairwise comparisons would greatly inflate the probability of Type I errors. Therefore, following the observation of Figure 3A, which suggested a turning point on the second day, the trend was further analyzed with two one-way RM-ANOVAs to respectively check the performance change in the first two days and the performance from day 2 to day 7. The first RM-ANOVA showed a significant IOR change from day 1 ($35.46 \pm 6.40$ ms) to day 2 ($5.39 \pm 3.24$ ms), $F(1, 15) = 31.53$, $p < .0001$, partial $\eta^2 = 0.68$. The second RM-ANOVA found that IOR stopped changing after day 2, $F(5, 75) = 0.23$, $p = .95$, partial $\eta^2 = 0.02$. This was also confirmed by a Bayesian RM-ANOVA, which showed strong evidence for the absence of any difference after day 2, $BF_{10} = 0.05$. Then, a one-sample $t$-test was carried out to evaluate the existence of IOR from day 2. Given that the IOR did not change across days 2 to 7, this $t$-test assessed the average IOR across these days against zero. The result showed that the average IOR ($7.25 \pm 3.28$ ms) was significant, $t(15) = 2.21$, $p = .04$, Cohen's $d = 0.55$, but only anecdotally supported by a Bayesian $t$-test, $BF_{10} = 1.68$.

**Discussion**

The results show that, in the search target-absent condition, IOR decreased when visual search efficiency improved during the training. This follows our prediction and agrees with the foraging facilitator hypothesis because less attentional engagement was required when the visual search became easier (due to the training) so that inhibitory tagging was less generated. Despite the decreasing trend over the training days, IOR stopped changing and remained stable after the second training day. This suggests that, although IOR plays an important role in foraging behavior, search performance improvement over training does not always relate to the change of IOR. This is probably due to the involvement of other factors such as top-down goals, bottom-up salience and prior search history (see Wolfe, 2021 for a review).

Another possibility is that the discrepancy in the changes of search efficiency and IOR in fact suggests the absence of any relationship between the two in this experiment. One may argue that the decrease of IOR after the first training day was due to the familiarization of the probe task (or some other procedural factors changed by the training), whereas the increase of search efficiency was driven by some other variables. However, this explanation is unlikely because of two reasons. First, as the foraging facilitator hypothesis suggests, IOR is a crucial mechanism in inefficient search. An abundant amount of studies observed the IOR effect in inefficient search (e.g., Müller & von Mühlenen, 2000; Thomas et al., 2006; Thomas & Lleras, 2009), demonstrating that IOR does discourage attention away from inspected locations, thereby facilitating search. It is hard to imagine that in the current study the IOR decrease during the first day had nothing to do with the search efficiency improvement.

Second, the current research data do not agree with the notion that the IOR decrease in the current data was only driven by procedural training. The target-present data (top-right corner of Figure 3) showed relatively strong and varied IOR over training days, against the possibility of a sole explanation of procedural learning in the first training day. This is also consistent with Xu et al. (2016), who investigated the change of IOR across training days utilizing a cue-target paradigm. Despite that their task was easier to learn than the current task, and that their paradigm employed more training with 384 trials per condition (240 trials in the current experiment), the IOR in Xu et al. (2016; Figure 3) kept decreasing over the first four days. As such, it is more likely that the IOR decrease was due to the reduced attentional engagement that resulted from the improved search efficiency.

It is also worth noting that despite the training, it is likely that the visual search remained serial to the last training day. Firstly, the search RTs in the pre-designated serial search conditions were relatively long, with a limited set size of 8 in the current study. These RTs were close to or longer than the RTs in previous work on serial search (Takeda & Yagi, 2000; Wang et al., 2010). Secondly, these RTs were much longer than the RTs in the parallel search condition in the current study. Thirdly, different from the parallel search in the present study, there was always a significant RT difference between the target-present and target-absent conditions in the serial search trials across all the training days. Despite the possibility, it is unlikely that all these three observations were caused by varying intercepts of the search slope function.

In Experiment 1, the search items were displayed as two clusters in two quadrants. This differed from the stimulus layouts used in previous studies in which the search items were

distributed evenly over the display (e.g., Klein, 1988; Müller & von Mühlenen, 2000). It

could be argued that the current results were specific to the spatial layout used. Therefore,

Experiment 2 used a traditional random layout to confirm the findings of Experiment 1.


## Experiment 2

**Method**

A new set of fourteen volunteers (12 females and 2 males, aged $20.50 \pm 0.55$ years, all right-

handed) with normal or corrected-to-normal vision took part in Experiment 2 for 6

consecutive days. One participant was removed due to the absence of IOR in the target-absent

condition on the first day (see Supplementary Materials for results including all participants).

The procedure was identical to that of Experiment 1 with three exceptions. Firstly, as

illustrated in Figure 1C, the search items were presented randomly at 8 of 16 possible

locations defined by an imaginary $4 \times 4$ grid covering the center ($7° \times 7°$) of the screen, with

a further $\pm 0.6°$ random vertical and horizontal displacement. Secondly, the fixation cross was

not presented during the search display and the probe display. Thirdly, a Cedrus RB-X40

response box was used to achieve a better RT resolution.


*Data analysis*

For the search task, trials with incorrect responses to the search targets (1.38% of total trials)

and RT outliers (detected with the same criteria as in Experiment 1; 2.46% of total trials)

were removed from the analysis. For the probe detection task, only the probe-present trials

were analyzed. Trials without probe detection responses (0.04% of total trials) were excluded

from analysis, as were trials preceded by incorrect search responses (1.31% of total trials) and

trials in which the probe and the search target occupied the same location (2.36% of total

trials). Finally, outliers were determined with the same criteria as those in Experiment 1,

leading to the removal of 1.81% of the data. The statistical analysis protocol was identical to

that in Experiment 1.

**Results**

*Search performance*

Figures 2 and 3 (bottom rows) illustrate the search performance. A two-way RM-ANOVA on

search efficiencies was carried out with search target presence (present vs. absent) and

training days (days 1-6) as factors. The main effect of search target presence was significant,

$F(1, 12) = 198.98$, $p < .0001$, partial $\eta^2 = 0.94$, with higher search efficiencies found in the

target-present trials ($28.35 \pm 2.56$ items/sec) than in the target-absent trials ($13.33 \pm 1.65$

items/sec). The main effect of training days reached significance as well, $F(5, 60) = 42.40$ , $p_c$

$< .0001$, partial $\eta^2 = 0.78$, showing a steady increase in the search efficiencies over the

training days and a significant linear contrast, $F(1, 12) = 74.63$, $p < .0001$, partial $\eta^2 = 0.86$.

The search efficiency improved from $10.07 \pm 1.02$ items/sec on the first training day to 30.08

$\pm 2.92$ items/sec on the last training day. There was also a significant search target presence $\times$

training days interaction, $F(5, 60) = 10.44$, $p_c < .0001$, partial $\eta^2 = 0.47$. Linear regression

slopes over the training days were obtained for all participants in each target-presence

condition. A paired-samples *t*-test between the target-present and -absent conditions revealed

that these two slopes differed significantly, $t(12) = 6.11$, $p < .0001$, Cohen's $d = 1.69$, with a

steeper slope in the target-present ($5.05 \pm 0.50$ items/sec/day) condition than in the target-

absent ($2.64 \pm 0.48$ items/sec/day) condition.

*IOR effects*

The IOR effects as a function of the training days in the search target-absent and search

target-present trials are shown in Figure 3 (bottom rows). A RM-ANOVA on IOR with factors

of search target presence (present vs. absent) and training days (days 1-6) found a significant

main effect of training days, $F(5, 60) = 4.51$ , $p = .001$, partial $\eta^2 = 0.27$, but no significant

main effect of search target presence, $F(1, 12) = 2.15$, $p = .17$, partial $\eta^2 = 0.15$.

The search target presence × training days interaction was significant, $F(5, 60) = 3.47$, $p_c$

$= .03$, partial $\eta^2 = 0.22$. To follow up the interaction, two one-way RM-ANOVAs were

employed to investigate the training effects in the target-absent and -present conditions

respectively. Despite that there was no difference across training days in the target-present

condition, $F(5, 60) = .31$, $p = .90$, partial $\eta^2 = 0.03$, the IOR effect significantly decreased

across the training days in the target-absent condition, $F(5, 60) = 9.24$, $p < .0001$, partial $\eta^2 =$

0.44. This training effect received significant contribution from both linear [$F(1, 12) = 17.92$,

$p = .001$, partial $\eta^2 = 0.60$] and quadratic contrasts [$F(1, 12) = 9.43$, $p = .01$, partial $\eta^2 = 0.44$].

The LMM analysis on IOR confirmed a significant exponential change of IOR over days,

$\chi^2(1) = 16.42$, $p < .0001$.

Similar to Experiment 1, further analyses were carried out to check these contrasts. First,

two one-way RM-ANOVAs checked the IOR decrease from day 1 to day 2, and the IOR

decreases after day 2. The results showed a significant decrease of IOR from day 1 ($16.67 \pm$

2.94 ms) to day 2 (0.59 ± 3.18 ms), $F(1, 12) = 25.70$, $p < .001$, partial $\eta^2 = 0.68$, but no significant change after that, $F(4, 48) = 1.03$, $p = .40$, partial $\eta^2 = 0.08$. The absence of IOR change after day 2 was also moderately supported by a Bayesian RM-ANOVA, $BF_{10} = 0.25$. The one-sample $t$-test on average IOR effects (−0.11 ± 1.23 ms) from days 2 to 6 showed the absence of IOR from day 2, $t(12) = .09$, $p = .93$, Cohen's $d = 0.03$, which was also moderately supported by Bayesian analysis, $BF_{10} = 0.28$.

**Discussion**

Experiment 2 replicated the results of Experiment 1 with a traditional search layout. This suggests that the current findings were not specific to a particular display layout. The two experiments convergingly illustrated that IOR decreased from Day 1 to Day 2 in the search target-absent condition while the search efficiency kept improving over the training days. This is consistent with the notion that when search efficiency improves, the attentional engagement becomes less, thus weaker inhibitory tags are generated.

Like in Experiment 1, it is likely that the visual search remained serial throughout all the training days because search RTs were always longer in the difficult inefficient search than in the parallel search and the inefficient search RTs were always longer in the search target-absent condition than in the target-present condition.

A key difference between the findings of the two experiments is that, from the second training day, IOR ceased to exist in Experiment 2 while it persisted at a low level throughout remaining days in Experiment 1. This is very likely due to the usage of the traditional layout, while the precise reason remains unclear. Nevertheless, the current findings might suggest

that search efficiency improvement sometimes is unrelated to the change of IOR. In certain situations, search efficiency improvement may not even require the existence of IOR after sufficient training. Given that inhibitory tagging is considered important in facilitating foraging behavior, this suggestion is intriguing.

## General discussion

The foraging facilitator hypothesis theorizes inhibitory tagging at previously inspected locations as a mechanism to optimize searching performance by prioritizing novel locations during visual search. The more challenging a visual search task is, the more serial attentional deployment is involved, yielding stronger inhibitory tagging (i.e., stronger IOR). Consistent with the foraging facilitator hypothesis, previous studies found that the probe detection RTs at previously attended locations were longer than those at new locations (i.e., the presence of the on-probe cost/IOR) and that IOR was stronger in serial than parallel search (Klein, 1988; Müller & von Mühlenen, 2000; Takeda & Yagi, 2000; Wang et al., 2010). The current study went beyond the comparison between these two search types and depicted how IOR (the measure of inhibitory tagging) changes when search efficiency (the indicator of search difficulty) gradually improved. Without changing the physical stimuli, task difficulty was gradually reduced with the long-term training protocols. Both experiments observed IOR on the first training day, replicating previous findings (e.g., Klein, 1988; Müller & von Mühlenen, 2000; Takeda & Yagi, 2000). More importantly, it was consistently found that, while search efficiency kept improving across training days, the IOR magnitude decreased until the second training day but ceased to change on the following days. The absence of IOR

changes from the second training day was not expected. This is an intriguing finding showing that search performance improvement does not always lead to the change of IOR. Fully understanding this effect requires further research. Nevertheless, the negative relationship[1] between search efficiency and IOR observed in both experiments supports the foraging facilitator hypothesis that IOR facilitates search performance via stronger inhibitory tagging in more difficult visual search.

One may argue that the current training protocol could have affected IOR through the procedural training in the probe task, independent from the search training, so that the search efficiency data and the IOR data were not related. From the theoretical perspective, although that the two tasks could potentially be separately trained, the performance in the probe task (i.e., IOR) is closely related to the performance in the search task (i.e., search efficiency) because the probe task is designed to disclose the attention deployment during visual search at the time of probe onset. Furthermore, despite that probe detection RTs decreased over the training days (Figure 2), it is unlikely that the IOR quickly diminished solely due to the procedural learning in the probe task, given that IOR in the search target-present condition in Experiment 1 (Figure 3, top-right corner) showed IOR (and variability in IOR) over all training days. Therefore, we believe that the relationship pattern between search efficiency and IOR found in the current study is intriguing and worth further research. To address the training effect on IOR, future studies could reduce or remove the probe task trials between the starting and ending days of the training.

---

[1] Consistent with this, an exploratory correlational analysis combining the two experiments' data showed a trend of negative relationship between search efficiency and IOR in the search target-absent condition on the first training day (see Supplementary Materials). This contradicts the prediction of a positive relationship by Klein and Dukewich (2006).

It is worth noting that the current research did not use the slope of the set-size function in the serial search condition to assess search efficiency, as was usually reported in the visual search literature (e.g., Klein, 1988), because only a single set size (8 items) was used in the current study. Instead, the search efficiency was quantified by dividing the difference in RTs between the serial search and the parallel search conditions by the number of search items. In the current approach, parallel search was used as a control condition to approximate the search efficiency for serial search. However, this approach might not be as accurate as the slope method. Furthermore, the current approach also raised the possibility that the gradual decrease in the serial search RTs was partially caused by the reduction of the intercept instead of the slope[2]. Further studies are needed to separate the slope decrease and the intercept decrease and to corroborate the approach adopted in the current study.

Unexpectedly, a discrepancy between search efficiency and IOR was found. Search efficiencies continued improving over the training days in both experiments, whereas IOR only decreased from Day 1 to Day 2 and remained unchanged since Day 2. This discrepancy might suggest that, unlike search efficiency, which has more capacity for improvement, IOR changes could be more limited and more easily reaching the floor. Indeed, after training, the target-distractor similarity in perception was reduced due to a heightened perceptual saliency of the target via training (Qu et al., 2017). With a higher perceptual saliency, the target tended to pop out more along with the training, leading to a shorter dwell time for each item (Becker, 2011), reduced inhibitory tagging (i.e., reduced IOR), and enhanced search efficiency. However, IOR is not the only factor contributing to search efficiency improvement. Visual

---

[2] We would like to thank Prof. Klein for raising this possibility.

search is guided by many factors such as top-down goals, bottom-up salience, and prior scene/search history (see Wolfe, 2021 for a review) in a winner-take-all fashion via a priority map where the contribution weights of these factors are flexibly adjusted (Fecteau & Munoz, 2006; Zelinsky & Bisley, 2015). In the current study, apparently IOR was not the main contributor to the performance improvement from the second training day. Visual search involves many other cognitive processes, such as sensory processing, attentional orienting, target discrimination, response preparation and selection (Nakayama & Martini, 2011), all of which could be enhanced during training (Clark et al., 2015). In the current study, it is likely that some or all of these processes kept improving after IOR reached the floor, and interestingly, maintained the linear trend of search efficiency improvement.

The data also showed a difference in IOR magnitudes between the two experiments. The IORs were stronger in Experiment 1 than Experiment 2 (not statistically assessed because it was not part of the research questions). Particularly, while IOR was present even after hitting the floor (day 2) in Experiment 1, IOR was absent since day 2 in Experiment 2. This agrees with our reasoning in the design phase, that the four-quadrant display layout should be able to generate stronger IOR than the traditional display layout. More importantly, Experiment 2 yielded higher search efficiency and weaker IOR than Experiment 1, again consistent with our hypothesis. This between-experiment difference might be due to the fast decay of the inhibitory tags. The weaker the inhibitory tags, the shorter the dwell time of the inhibitory tags. If this is plausible, in Experiment 2, the inhibitory tags might have diminished when the experiment measured the IOR with the probe. This hypothesis requires further research. Nevertheless, the results from Experiment 2 clearly showed that, in contrast to common

belief, serial search does not necessarily generate inhibitory tags. It is speculated that the inhibitory tagging process is only active when the serial search task's difficulty is above a threshold, with some variability across different tasks. When the search is not highly demanding, inhibitory tagging may not be fully engaged or may decay fast.

The absence of IOR from Day 2 in Experiment 2 may also be partially attributed to the probability of target presence in the post-search probe detection task. In the current study, the probe was present in 80% of trials, rather than 50% as in Klein (1988). Given that to detect the probe means to attend to previously searched locations, the high probability of probe appearance in the current study could encourage attention returning to the previously inspected locations, thus counteracting the IOR effect. This suggestion is supported by the finding that IOR was eliminated when the target appeared more frequently at a previously inspected location (Farrell et al., 2010). This counteracting effect might be less an issue on the first training day, because the search task was demanding and yielding a strong IOR. However, after the first training day, the participants became more proficient with the task, therefore could possibly afford to divert more resource from the search task to the probe detection task and allow more attentional revisiting of the searched locations due to the higher prevalence of probes appearing at the inspected locations.

Notwithstanding the discussion of various possibilities above (which will require further investigation), the current results made an important addition to the literature on the foraging facilitator hypothesis by suggesting that inhibitory tagging (as a result of attentional engagement) is an adaptive mechanism of the visual system. Specifically, the magnitude of IOR was reduced, and sometimes fully diminished, when search efficiency improved

(weakening attentional engagement) during the training. Under a high search difficulty, when the inhibitory tags could significantly improve search efficiency, they were employed; under a decreased search difficulty, when there was little advantage in employing them, they did not fully engage or decayed quickly. Overall, we have shown how IOR varied during long-term training that improved search efficiency. The current findings support the foraging facilitator hypothesis that IOR improves search efficiency via stronger inhibitory tagging in more difficult search.

**Open practices statement:** All experiments data and analysis materials will be available at https://github.com/yangzhangpsy/trainingOnVsAndIOR, and none of the experiments was preregistered.

# References

Becker, S. I. (2011). Determinants of dwell time in visual search: similarity or perceptual difficulty? *PLoS One*, *6*(3), e17740. https://doi.org/10.1371/journal.pone.0017740

Bennett, P. J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological science*, *12*(1), 76-80. https://doi.org/10.1111/1467-9280.00313

Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods*, *44*(4), 1255-1265.

https://doi.org/10.3758/s13428-012-0186-0

Chun, M. M., & Wolfe, J. M. (1996). Just say no: how are visual searches terminated when there is no target present? *Cognitive Psychology*, *30*(1), 39-78. https://doi.org/10.1006/cogp.1996.0002

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

Farrell, S., Ludwig, C. J. H., Ellis, L. A., & Gilchrist, I. D. (2010). Influence of environmental statistics on inhibition of saccadic return. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(2), 929-934. https://doi.org/10.1073/pnas.0906845107

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382-390. https://doi.org/10.1016/j.tics.2006.06.011

Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, *10*(19), 1209-1212. https://www.cell.com/current-biology/pdf/S0960-9822(00)00729-6.pdf

Klein, R. M. (1988). Inhibitory tagging system facilitates visual search. *Nature*, *334*(6181), 430-431. https://doi.org/10.1038/334430a0

Klein, R. M., & Dukewich, K. (2006). Does the inspector have a memory? *Visual Cognition*, *14*(4-8), 648-667. https://doi.org/10.1080/13506280500194022

Klein, R. M., & Taylor, T. L. (1994). Categories of cognitive inhibition with reference to attention. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 113-150). New York: Academic Press.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 14. https://doi.org/10.1068/v070821

Li, A. S., Miao, C. G., Han, Y., He, X., & Zhang, Y. (2018). Electrophysiological correlates of the effect of task difficulty on inhibition of return. *Frontiers in Psychology*, *9*, 2403. https://doi.org/10.3389/fpsyg.2018.02403

Li, A. S., Zhang, G. L., Miao, C. G., Wang, S., Zhang, M., & Zhang, Y. (2017). The time course of inhibition of return: Evidence from Steady-State Visual Evoked Potentials. *Frontiers in Psychology*, *8*, 1562. https://doi.org/10.3389/fpsyg.2017.01562

Lin, Z., Miao, C., & Zhang, Y. (2020). Human electrophysiology reveals delayed but enhanced selection in inhibition of return. *Cognition*, *205*, 104462. https://doi.org/10.1016/j.cognition.2020.104462

Ma, W. J., Navalpakkam, V., Beck, J. M., Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nature Neuroscience*, *14*(6), 783-790. https://doi.org/10.1038/nn.2814

Müller, H. J., & von Mühlenen, A. (2000). Probing distractor inhibition in visual search: inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1591-1605. https://doi.org/10.1037//0096-1523.26.5.1591

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387-391. https://doi.org/10.1038/nature03390

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*(13), 1526-1537. https://doi.org/10.1016/j.visres.2010.09.003

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bowhuis (Eds.),

*Attention and performance X* (pp. 531-556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive neuropsychology*, *2*(3), 211-228. https://doi.org/10.1080/02643298508252866

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

Takeda, Y., & Yagi, A. (2000). Inhibitory tagging in visual search can be found if search stimuli remain visible. *Perception and Psychophysics*, *62*(5), 927-934. https://doi.org/10.3758/bf03212078

Thomas, L. E., Ambinder, M. S., Hsieh, B., Levinthal, B., Crowell, J. A., Irwin, D. E., Kramer, A. F., Lleras, A., Simons, D. J., & Wang, R. F. (2006). Fruitful visual search: Inhibition of return in a virtual foraging task. *Psychonomic Bulletin and Review*, *13*(5), 891-895. https://doi.org/10.3758/bf03194015

Thomas, L. E., & Lleras, A. (2009). Inhibitory tagging in an interrupted visual search. *Attention, Perception, & Psychophysics*, *71*(6), 1241-1250. https://doi.org/10.3758/app.71.6.1241

Vanselst, M., & Jolicoeur, P. (1994). A solution to the effect of sample-size on outlier elimination. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *47A*(3), 631-650. https://doi.org/10.1080/14640749408401131

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research*, *50*(2), 220-228. https://doi.org/10.1016/j.visres.2009.11.013

Wang, Z., Zhang, K., & Klein, R. M. (2010). Inhibition of return in static but not necessarily in dynamic search. *Attention, Perception and Psychophysics*, *72*(1), 76-85. https://doi.org/10.3758/app.72.1.76

West, B. T., Welch, K. B., & Galecki, A. T. (2014). *Linear mixed models: a practical guide using statistical software* (2nd ed.). Crc Press.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences*, *7*(2), 70-76. https://doi.org/10.1016/s1364-6613(02)00024-4

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin and Review*, *28*(4), 1060-1092. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., & Pokorny, C. W. (1990). Inhibitory tagging in visual search: A failure to replicate. *Perception and Psychophysics*, *48*(4), 357-362. https://doi.org/10.3758/bf03206686

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

Zhang, F., Lin, Z., Zhang, Y., & Zhang, M. (2021). Behavioral evidence for attention selection as entrained synchronization without awareness. *Journal of Experimental Psychology: General*, *150*(9), 1710-1721. https://doi.org/10.1037/xge0000825

Zhang, G. L., Li, A. S., Miao, C. G., He, X., Zhang, M., & Zhang, Y. (2018). A consumer-grade LCD monitor for precise visual stimulation. *Behavior Research Methods*, *50*(4), 1496-1502. https://doi.org/10.3758/s13428-018-1018-7