Flanker tasks based on congruency manipulation are biased measures of selective attention in perceptual load studies

Zhi Li[1] · Jiafei Lou[1]

Abstract
Flanker tasks based on a flanker–target congruency manipulation are widely used in perceptual load studies to investigate under what circumstances task-irrelevant flankers may be processed. An implicit assumption underlying the congruency manipulation is that the three types of flankers (congruent, incongruent, and neutral) attract attention homogeneously. However, in the present study, we provide evidence to demonstrate that this assumption is wrong: We discovered that incongruent/congruent flankers attracted more attention than the neutral flanker did. To avoid this attentional bias induced by the flanker–target congruency manipulation, we developed a new flanker paradigm in which the extent of flanker processing was evaluated by comparing the threshold stimulus exposure durations (TSEDs) for successfully performing a target identification task when a task-irrelevant flanker was presented versus when the flanker was absent. The flanker was processed if the TSED was longer when the flanker was present than when it was absent. This new paradigm provides an unbiased measure of selective attention when neutral flankers are used. The present data, obtained with neutral flankers in the new paradigm, were consistent with the dilution theory of selective attention, but inconsistent with the perceptual load theory of selective attention.

Keywords Selective attention · Top-down control · Load theory · Dilution theory · Activity distribution model

Early theories of selective attention mostly focused on when the stimulus information is attended to (Broadbent, 1958; Deutsch & Deutsch, 1963; Treisman & Gelade, 1980), whereas recent theories of selective attention have been more concerned with how attentional resources are divided up among competing sources of information. The perceptual load theory of selective attention hypothesizes that attention is allocated first to the task-relevant information. Only when the task-relevant information is fully processed and extra attentional resources are left do the unconsumed resources spill over to the task-irrelevant information (Lavie & Cox, 1997; Lavie & Tsal, 1994). The idea of load theory has inspired a large number of studies to test its validity. Although both supporting evidence (Forster & Lavie, 2007, 2008; Lavie, 2005; Lavie & de Fockert, 2003) and contradicting findings (Eltiti, Wallace, & Fox, 2005; Johnson, McGrath, & McNeil, 2002;

Theeuwes, Kramer, & Belopolsky, 2004; Tsal & Benoni, 2010; Wilson, Muroi, & MacLeod, 2011) have been reported, most perceptual load studies have one thing in common: they have used some form of flanker task to study the circumstances under which the task-irrelevant flanker is processed. In the classic flanker task, the target is always presented at the fixation point (Eriksen & Eriksen, 1974). But in the flanker tasks used in load studies, observers have to search for the target, and they know in advance where a flanker might appear. The question is whether and when observers inhibit the identification of the flanker by using their preknowledge of the flanker's potential locations. The target usually has two alternative identities, and the task is to determine which identity is presented in each trial. The flankers have either the target identity (congruent flankers), the alternative identity for the target (incongruent flankers), or neither (neutral flankers). The congruency effect (the common result that correct response times [RTs] are longer in the incongruent than in the congruent or neutral condition) is used to evaluate the flanker processing.

An implicit assumption when using the congruency effect to evaluate flanker processing is that the three types of flankers (i.e., congruent, incongruent, and neutral) are homogeneous in

✉ Zhi Li
zli1@zju.edu.cn

[1] Department of Psychology and Behavioral Sciences, Zhejiang University (Xixi Campus), Hangzhou, China

attracting attention. However, this assumption might be wrong. Models of attention such as the feature integration theory (Treisman & Gelade, 1980), guided search (Wolfe, 1994), and the biased-competition model (Desimone & Duncan, 1995) have all proposed that the deployment of attention is an interactive process between bottom-up information and top-down knowledge. Attention is guided by top-down knowledge about the target. In visual search, constantly changing the target from trial to trial (as opposed to using a fixed target) slows down the RT (Schneider & Shiffrin, 1977). There is evidence that a stimulus matching the target template is more attractive of attention than is a neutral stimulus; for example, increased similarity between the distractor and target reduces the efficiency of visual search (Duncan & Humphreys, 1989). Moreover, in the spatial-cueing paradigm, involuntary shifts of attention to a given stimulus event are contingent on whether that event shares a feature property that is critical to the performance of the task at hand (Folk, Remington, & Johnston, 1992). If the distractor–target similarity matters in the visual-search and spatial-cueing paradigms, then in the flanker paradigm, the congruent and incongruent flankers (which match the target) might also be more attractive of attention than the neutral flanker. The flanker tasks tested in load studies have often used letters as the visual stimuli. This raised the question of whether similarity between the letters might also affect the allocation of attention, because theories of attention often assume that complex features, such as letter identity, cannot guide attention (Krueger, 1984; Wolfe & Horowitz, 2017). We think the answer is affirmative, because the similarity effect might also occur at the level of a physical property (i.e., shape), which would not require semantic processing. Moreover, it has been shown that visual attention is sensitive even to the semantic properties of visual objects (Lupyan, 2008) and to alphanumeric category (Avital-Cohen & Tsal, 2016).

A cautious reader may note that the flanker paradigm is different from the visual-search and spatial-cueing paradigms in that, in the latter two paradigms, the target shares the same spatial locations with the distractors, whereas in the flanker paradigm, the distractor (i.e., the flanker) never appears in the potential target locations. That is, top-down spatial knowledge exists in the flanker tasks that biases attention away from the flanker. The existence of such a top-down spatial bias makes it unclear whether flanker–target similarity might affect the flanker processing in the flanker paradigm. A recent study has demonstrated that, in the flanker paradigm, the extent of flanker processing can be affected by top-down guidance (Avital-Cohen & Tsal, 2016). In the present study, we will further show that, in the flanker paradigm, flanker processing is not only subject to top-down guiding, but also affected by flanker–target similarity. More specifically, we will show that incongruent and congruent flankers are more attractive of attention than is a neutral flanker.

Experiment 1

A flanker paradigm without a congruency manipulation was used. Observers performed a target identification task. In the baseline condition, a letter array was shown. Observers searched for and identified the target letter in this array. In the flanker condition, a task-irrelevant flanker was also presented to the side of this search array. The threshold stimulus exposure durations (TSEDs) for successfully fulfilling this task were compared between the two conditions. A previous study had shown that such a target identification task might involve a serial scanning process, so that adding each extra letter to the search array increases the TSED by 40 ms (Li, Xin, Li, & Li, 2018). Thus, if the TSED in the flanker condition were significantly greater than the TSED in the baseline condition, this would be a clear sign that the flanker in the flanker condition was processed to some extent, and therefore delayed the TSED. Unlike the RT paradigm, the TSED paradigm does not require prompt responses. The observers are instead required to give the top priority to response accuracy and to take their time when making responses. Therefore, there is virtually no speed–accuracy trade-off in the TSED paradigm. This feature makes the TSED paradigm less affected by processes at the response-selection stage, and also makes it immune to the stimulus–response congruence effect seen in the RT paradigm.

Method

Observers A priori power analysis was conducted with the program G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). The repeated measures analysis of variance (ANOVA) with a moderate effect size of 0.25 (Cohen, 1988), $\alpha = .05$, and $1 - \beta = .80$ gave a statistical power of 81.98% and a sample size of a minimum of 16 observers. Considering publication bias and uncertainty (Anderson, Kelley, & Maxwell, 2017), 24 students (13 females, 11 males) from Zhejiang University participated in Experiment 1. The observers ranged from 18 to 21 years of age ($M = 19$). They all had normal or corrected-to-normal vision. The experimental procedures reported in this article were approved in advance by the local research ethics committee and were conducted in accordance with the Declaration of Helsinki.

Design The purpose of Experiment 1 was to investigate whether incongruent and congruent flankers are more attractive of attention than neutral flankers. Thus, three flanker conditions (incongruent, congruent, and neutral flankers) and a baseline condition (no flankers) were tested. Two display set sizes were included, to examine the possible effect of set size. Altogether, eight stimulus conditions—that is, (3 flanker conditions + 1 baseline condition) × 2 set sizes—were tested. The dependent variable was the TSED for successfully fulfilling a

target identification task. The baseline TSED was compared to the TSED in each of the three flanker conditions. A within-subjects design was used, with the eight stimulus condition trials being randomly mixed. If flanker–target similarity does not affect the extent of flanker processing, the differences in TSED between the baseline condition and the three flanker conditions should be identical. Otherwise, the results would support the hypothesis that the three different types of flankers attract attention inhomogeneously.

Stimuli The stimuli consisted of white letters appearing on a black background (Fig. 1). The target letter (either an H or S) appeared randomly in one of six locations evenly spaced on an imaginary circle centered at fixation. In the rest of the locations, either identical short lines (for the small set-size condition) or different letters (for the large set-size condition; the letters were chosen from among A, C, E, F, L, O, or P) were presented. Each letter subtended a visual angle of 0.72° in height and 0.4° in width. The center-to-center distance between letters on the imaginary circle was 1.24°. In the baseline conditions, observers determined whether H or S was shown in the circular array. In the flanker conditions, a larger flanker letter (0.96° × 0.53° in visual angle, chosen from among H, S, A, C, E, F, L, O, or P) appeared equally often to the left or the right side of the circular array. The center-to-center distance between the imaginary circle and the flanker was 3.2°. The observers were asked to determine whether H or S was presented in the circular array, while ignoring the flanker letter. No two letters were identical in the circular array. The stimuli were presented on a CRT monitor (SONY Co.) at a 160-Hz frame rate. Observers sit with their headrest on a chinrest. The viewing distance was 57 cm.
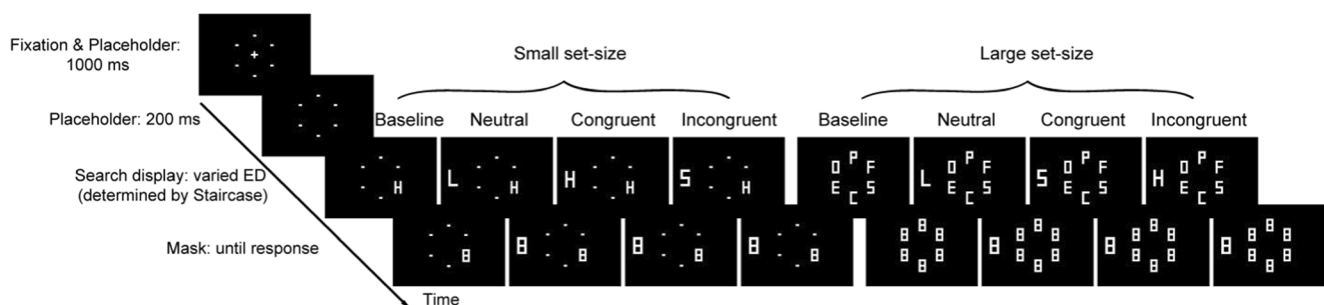
Procedure Each trial began with the presentation of a fixation cross and six short lines marking the possible target locations. The fixation cross disappeared after 1,000 ms, but the short lines were still presented for another 200 ms, which allowed attention to disengage from the fixation cross. The search display (i.e., the central letter array) appeared on the screen for a varied exposure duration, determined by a staircase procedure. The letters then turned into block figure-8s that served

as backward masks. Observers made a two-alternative forced choice response, in which they indicated whether an H or an S had been presented in the circular array, by pressing the "H" or "S" key on a keyboard. They were told that response accuracy should be given the top priority. An audible tone was played after each incorrect response. A 1-up, 3-down staircase procedure was used to obtain the TSED for successful searches in each of the eight stimulus conditions. The initial exposure duration (ED) was 87.5 ms for the small set-size conditions, and it was 300 ms for the large set-size conditions. This duration was decreased by 6.25 ms (i.e., the duration of a frame) after three consecutive correct responses, and it was increased by 6.25 ms after each incorrect response. A "turn" was defined when two consecutive changes made to the exposure duration differed (from a decrease change to an increase change, or vice versa). Ten turns were required for each staircase. The averaged ED over the last four turn points was used to determine the TSED. Each observer completed 32 practice trials at the beginning of the experiment. The exposure duration was fixed in the practice trials (i.e., 300 ms for the small set-size conditions, and 750 ms for the large set-size conditions). If the accuracy rate in the practice session was lower than 90%, the observer completed another 32 practice trials. Most observers passed this criterion in their first practice session. On average, about 477 trials were finished by each participant in Experiment 1, which lasted about 40 min.
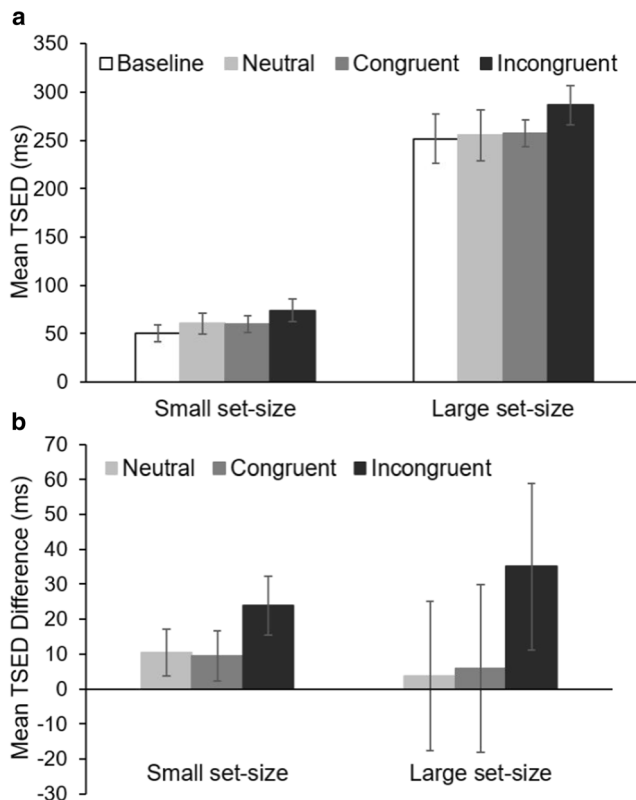
Data analyses The data were analyzed via repeated measures ANOVA. For all inferential statistics reported in the present article, alpha was set at .05. Where the sphericity assumption was violated, as measured with Mauchly's sphericity test, Greenhouse–Geisser-corrected values are reported.

Results

The mean TSEDs for successful searches in the eight stimulus conditions of Experiment 1 are plotted in Fig. 2a. As expected, the TSED in the baseline condition was the shortest among all testing conditions, for both small and large set sizes. We cared more about the TSED difference between each flanker type and its

Fig. 1 Paradigm used in Experiment 1. The eight stimulus conditions are plotted in separate columns. Observers judged whether a letter H or S was presented in the central circular array

Fig. 2 Results of Experiment 1. (a) Mean threshold stimulus exposure durations (TSEDs) for successfully fulfilling the target identification task in the eight stimulus conditions. (b) Mean TSED difference between the baseline and each flanker condition. Error bars indicate the 95% confidence intervals

.05), whereas for the large set size, the difference was significantly greater than zero only for the incongruent flanker ($p < .01$). The mean response accuracy rate for trials after the fifth turn in the staircase across the eight testing conditions was 81.2%, suggesting that the 1-up, 3-down staircase (with a nominal accuracy rate of 79.4%) worked as expected.

Discussion

The crucial finding of Experiment 1 was that the TSED differences between the baseline and the three flanker conditions were not identical, which demonstrated that the three types of flankers attracted attention inhomogeneously. Although the result that the TSED difference for incongruent flankers was greater than those for congruent/neutral flankers is apparently similar to the typical congruency effect, we must keep in mind that the TSED paradigm used in this experiment did not rely on a congruency manipulation. The fact that the TSED difference for congruent flankers was similar to that for neutral flankers might have been due to the following reason: If the TSED in the baseline condition was T and the ED was set to T for the neutral flanker condition, because the neutral flanker was less attractive of attention, an ED of time T would be just enough to successfully fulfill the task. Therefore, the TSED for the neutral flanker might have been similar to that for the baseline. If the baseline TSED was T and the ED was set to T for the congruent flanker condition, with a certain probability, the congruent flanker might be identified first (due to its attractiveness), so that there would not be enough time to identify the target. However, observers could still guess the identity of the target, and this guessing response could be implicitly influenced by the identified identity of the congruent flanker, because it was a candidate identity of the target. That is, the guessing response might be biased toward the correct answer, which would make the TSED for the congruent flanker appear smaller than it is. Similarly, the incongruent flanker might also be identified first and require the observer to guess the target identity. But this time, the guessing response would be biased toward the incorrect answer, making the TSED for the incongruent flanker appear larger than it should be. Nonetheless, the underestimation in the TSED of the congruent flanker and the overestimation in the TSED of the incongruent flanker would be built on the fact that the two types of flankers attract attention. Although the neutral flanker attracts attention when the set size is small, it does not attract attention when the set size is large. Thus, at least in the large set-size condition, it is clear that the incongruent flanker attracts more attention than the neutral flanker does, which is inconsistent with the idea that the three types of flankers attract attention homogenously.

baseline, because this reflected the absolute amount of attention allocated to that type of flanker. If the TSED difference between a flanker type and its baseline were not different from zero, this would suggest that no attention was allocated to that type of flanker. In contrast, if the TSED difference were significantly greater than zero, this would suggest that a certain amount of attention was allocated to that type of flanker. For this reason, the mean TSED difference between the baseline and each type of flanker condition is plotted in Fig. 2b. A two-way (3 flanker types × 2 set sizes) ANOVA revealed a significant main effect of flanker type, $F(2, 46) = 6.785$, $p < .01$, $\eta_p^2 = .228$, but not of set size, $F(1, 23) = 0.001$, $p = .975$, $\eta_p^2 < .001$, and the interaction between flanker type and set size was not significant, $F(2, 46) = 1.396$, $p = .258$, $\eta_p^2 = .057$. Bonferroni post-hoc tests further revealed that the TSED difference between the baseline and flanker conditions was significantly larger for incongruent flankers than for either congruent flankers, $p = .022$, or neutral flankers, $p < .01$. However, the TSED difference for the congruent flanker was not different from that for the neutral flanker, $p = 1.0$. One-sample t tests further revealed that, for the small set size, the TSED difference was significantly greater than zero for all flanker types (all ps <
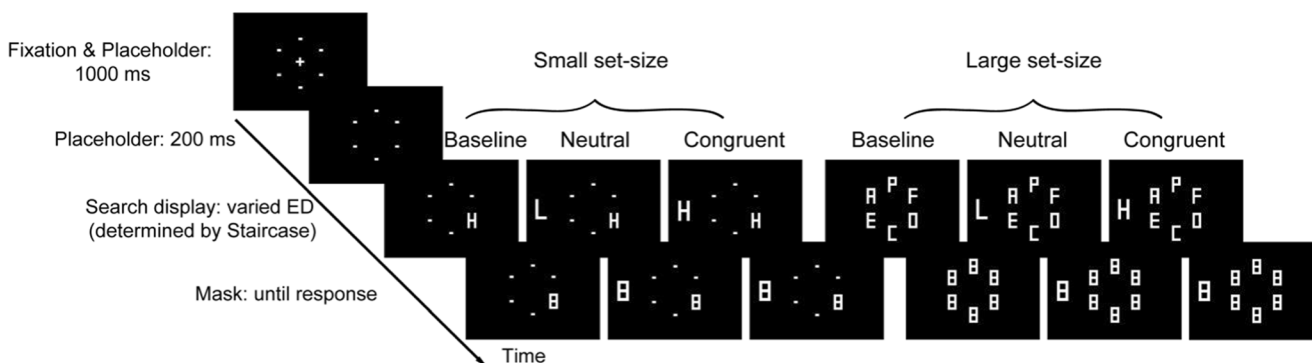
Experiment 2a

Experiment 2a was conducted to show that when catching sight of the identity of the congruent flanker does not help the guessing response, the TSED for the congruent flanker will exceed the TSED for the neutral flanker. In Experiment 2a, the target had a single identity, and the flanker was either the same as the target (the congruent flanker) or different from the target (the neutral flanker). The target was presented in half the trials, and observers judged whether the target was presented.

Method

Observers An a priori power analysis was conducted with the program G*Power. A repeated measures ANOVA with an effect size of 0.25, $\alpha = .05$, and $1 - \beta = .80$ gave a statistical power of 81.98% and a sample size of a minimum of 19 observers. Considering publication bias and uncertainty, 24 students (12 females, 12 males) from Zhejiang University were recruited in Experiment 2a. The observers ranged from 18 to 27 years of age ($M = 21$), with normal or corrected-to-normal vision. None had participated in Experiment 1.

Design, stimuli, and procedure The design, stimuli, and procedure used in Experiment 2a were identical to those used in Experiment 1, except for a few changes (Fig. 3). First, the target had only a single identity—that is, observers were asked to determine whether a letter H was presented in the central array. They responded by pressing the left-arrow key (H absent) or the right-arrow key (H present). The H was presented only in half the trials. Second, only six stimulus conditions were tested; that is, the baseline condition and two types of flanker conditions (neutral and congruent) were paired with the two set sizes (small and large). The neutral flanker was chosen from among S, A, C, E, F, L, O, and P, and the congruent flanker was H. On average, about 364 trials were finished by each participant in Experiment 2a, which lasted about 30 min.

Results

The mean TSEDs of the successful searches in the six testing conditions of Experiment 2a are plotted in Fig. 4a. Because we cared about whether the extent of flanker processing was greater for the congruent flanker than for the neutral flanker, the mean TSED difference between the baseline and each flanker condition is plotted in Fig. 4b. A two-way (2 flanker types × 2 set sizes) ANOVA was conducted on the difference data. A significant main effect of flanker type was observed, $F(1, 23) = 7.83$, $p = .01$, $\eta_p^2 = .254$, but no significant main effect of set size, $F(1, 23) = .297$, $p = .591$, $\eta_p^2 = .013$, nor an interaction between flanker type and set size, $F(1, 23) = 1.776$, $p = .196$, $\eta_p^2 = .072$. One-sample t tests further showed that, for the small set size, the TSED difference was significantly greater than zero for both the congruent and neutral flankers (all ps $< .05$), but for the large set size, the TSED difference was significantly greater than zero only for the incongruent flanker ($p = .032$). The averaged response accuracy rate for trials after the fifth turn in the staircase procedure across the six testing conditions was 80.0%, consistent with the expected accuracy rate.

Discussion

The significant main effect of flanker type in Experiment 2a shows that the extent of flanker processing for the congruent flanker was greater than that for the neutral flanker, which suggests that the congruent flanker is more attractive of attention than the neutral flanker. In fact, in Experiment 2a the identity of the congruent flanker still biased the responses, similar to what we observed in Experiment 1. The congruent flanker attracted more attention than the neutral flanker did, which would have two consequences. First, more time would be required to determine whether an H was presented in the central array, which would increase the TSED of the congruent flanker. Second, under short exposure durations, there would be no extra time to decide whether an H was presented in the central array. Observers would have to guess. They
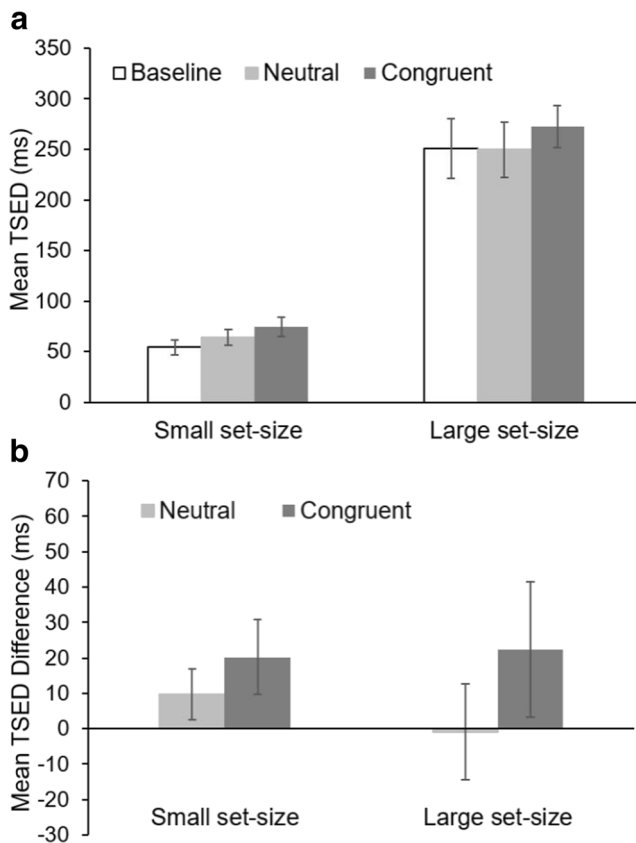

Fig. 3 Paradigm used in Experiment 2a. The six stimulus conditions are plotted in separate columns. Observers determined whether the letter H was presented in the central circular array

Fig. 4 Results of Experiment 2a. (a) Mean threshold stimulus exposure durations (TSEDs) for successfully fulfilling the target detection task in the six stimulus conditions. (b) Mean TSED difference between the baseline and each flanker condition. Error bars indicate the 95% confidence intervals

Experiments 1 and 2a adopted different strategies. First, in Experiment 2a, knowing whether an H was shown in the flanker location would not help in the task (i.e., reporting whether an H was presented in the target location). The strategy of first searching for an H and then determining whether it was in the target location would be very inefficient in Experiment 2a. We doubt that observers would use such an inefficient strategy, although observers might not always adopt the most efficient strategy in visual search (Chen & Cave, 2016). Another more powerful reason to doubt that different strategies would be adopted in Experiments 1 and 2a actually comes from the TSED data of the neutral flanker conditions in the two experiments, where the neutral flanker was less attractive of attention. If the observers in Experiment 1 focused on the target location, the strength of flanker processing (the TSED difference between the flanker and baseline conditions) should have been fairly small for the neutral flankers. In contrast, if the observers in Experiment 2a searched both the flanker location and target location, the strength of flanker processing for the neutral flankers should have been relatively larger than in Experiment 1. However, the data showed this was not the case. Processing for the neutral flankers in Experiment 1 (i.e., small set size: $M = 10.4$ ms, 95% CI = 6.7 ms; large set size: $M = 3.8$ ms, 95% CI = 21.4 ms) was not significantly different from that in Experiment 2a (i.e., small set size: $M = 9.7$ ms, 95% CI = 7.2 ms; large set size: $M = -0.8$ ms, 95% CI = 13.6 ms), $F(1, 46) = 0.150$, $p = .700$, $\eta_p^2 = .003$. Therefore, we believe that the observers in Experiments 1 and 2a did not adopt different search strategies.

Experiment 2b

A potential problem with Experiment 2a was that the neutral trials consisted of eight flanker letters (S, A, C, E, F, L, O, and P), but the congruent trials only of Hs. Thus, Hs occurred as flankers eight times as often. It has been shown that stimulus contingency is a cue that affects the allocation of attention (Melara & Algom, 2003). Perhaps simply seeing the Hs eight times as often would be sufficient to produce the effects in Experiment 2a. To rule out this possibility, we conducted Experiment 2b, in which the neutral trials consisted of only one flanker letter.

Method

Observers An a priori power analysis was conducted with the program G*Power. A repeated measures ANOVA with an effect size of 0.25, $\alpha = .05$, and $1 - \beta = .80$ gave a statistical power of 81.98% and a sample size of a minimum of 19 observers. Considering publication bias and uncertainty, 24 students (15 females, nine males) from Zhejiang University were recruited in Experiment 2b. The observers ranged from

would be more likely to guess that the H was presented if they saw an H in the flanker location. But this guess would cause a wrong response in 50% of the cases, which should also increase the TSED for the congruent flanker. Consistent with this idea, in Experiment 2a the TSED for the congruent flanker was significantly greater than that for the neutral flanker.

It worth noting that the task used in Experiment 2a was single-target search, but the task in Experiment 1, similar to most flanker tasks used in load studies, was dual-target search. It has been reported that observers are reluctant to guide search with two different targets (Menneer, Barrett, Phillips, Donnelly, & Cave, 2007; Menneer, Cave, & Donnelly, 2009). Perhaps the observers in Experiment 2a adopted a search strategy different from that used in Experiment 1. That is, in Experiment 1 the observers might have focused their attention on the target location and determined which target was present, whereas in Experiment 2a they might have first searched for an H and then decided whether it was in a target or a flanker location. If that was the case, it would make the conclusion from Experiment 2a not generalizable to the flanker tasks used in load studies. However, there are good reasons to reject the possibility that the observers in

19 to 29 years of age ($M = 23$), with normal or corrected-to-normal vision. None had participated in Experiment 1 or 2a.

Design, stimuli, and procedure The design, stimuli, and procedure used in Experiment 2b were identical to those used in Experiment 2a, except that the neutral trials consisted of only the letter L. The congruent trials, neutral trials, and baseline trials were randomly mixed. The neutral flanker was always L, and the congruent flanker was always H. On average, 376 trials were finished by each participant in Experiment 2b, which lasted about 30 min.

Results

The mean TSEDs of the successful searches in the six testing conditions of Experiment 2b are shown in Fig. 5a. The mean TSED difference between the baseline and each flanker condition is plotted in Fig. 5b. A two-way (2 flanker types × 2 set sizes) ANOVA was conducted on the difference data, where a significant main effect of flanker type was observed, $F(1, 23) = 11.55$, $p = .002$, $\eta_p^2 = .334$, but no significant main effect of set size, $F(1, 23) = 0.949$, $p = .340$, $\eta_p^2 = .040$, or interaction between flanker type and set size, $F(1, 23) = 0.782$, $p = .386$,
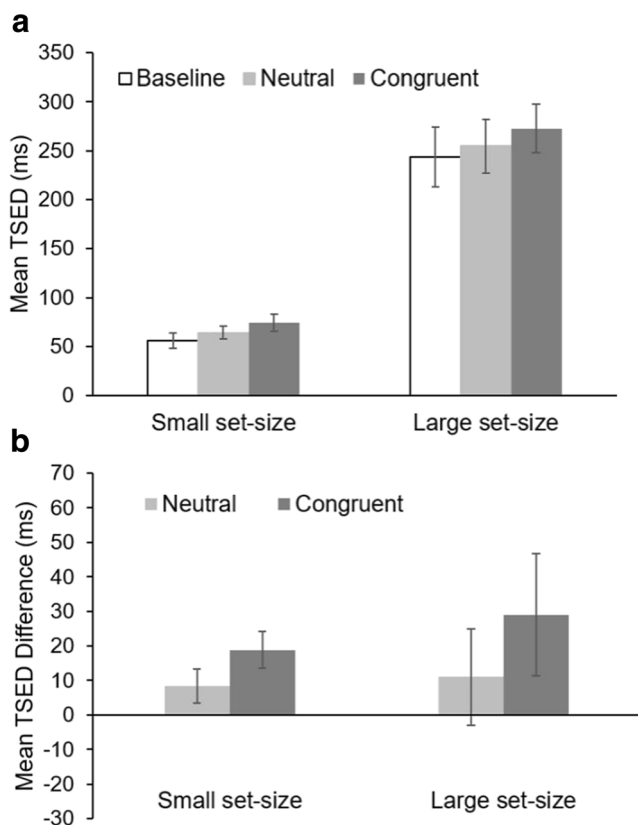
Fig. 5 Results of Experiment 2b. (a) Mean threshold stimulus exposure durations (TSEDs) for successfully fulfilling the target detection task in the six stimulus conditions. (b) Mean TSED difference between the baseline and each flanker condition. Error bars indicate the 95% confidence intervals

$\eta_p^2 = .033$. One-sample t tests further showed that, for the small set size, the TSED difference was significantly greater than zero for both the congruent and neutral flankers (all ps < .01), but for the large set size, the TSED difference was significantly greater than zero only for the incongruent flanker ($p = .004$). The mean response accuracy rate on trials after the fifth turn in the staircase procedure across all testing conditions was 81.8%, as expected.

Discussion

The results of Experiment 2b well replicated the findings of Experiment 2a, which ruled out the possibility that the TSED difference between the neutral and congruent trials was due to stimulus contingency.

Experiment 3

It has been known for some time that the flanker effect in the flanker task may have two components: the stimulus–stimulus congruence effect and the stimulus–response compatibility effect (van Veen, Cohen, Botvinick, Stenger, & Carter, 2001). The stimulus–stimulus congruence effect is assumed to occur at the stimulus identification level, whereas the stimulus–response congruence effect is thought to occur at the response selection stage. The effects of these two components are additive (Scerrati, Lugli, Nicoletti, & Umiltà, 2017). In the present study, the limiting factor was pushed toward encoding and discrimination of the target, so the stimulus–response congruence effect cannot be responsible for the difference in TSED differences across flanker types. However, in the small set-size conditions of Experiments 1 and 2, the flanker was processed irrespective of its flanker type (i.e., the TSED difference was always greater than zero), which opens the possibility that the difference in TSED differences between the three types of flankers in the small set-size condition was due to the stimulus–stimulus congruence effect rather than the inhomogeneous attractiveness in the different types of flankers. That is, it is plausible that the three types of flankers could have attracted attention homogeneously and caused reliable but identical TSED differences if, on top of that, an additional stimulus–stimulus congruence effect caused the difference (e.g., between neutral and incongruent flankers) in TSED differences. Experiment 3 investigated whether inhomogeneous attractiveness was actually responsible for the difference in TSED differences.

It has long been known that abrupt stimulus onsets are more salient than gradual stimulus onsets (or offset stimuli). It has been demonstrated using the spatial-cueing paradigm that abrupt stimulus onsets capture attention obligatorily, whereas stimulus offsets do not (Wright & Ward, 2008). The reason the neutral flankers were

somehow processed in the small set-size condition of Experiments 1 and 2 might have been because the abrupt stimulus onset attracted attention. Therefore, offset stimuli were used in Experiment 3. We hypothesized that when using the offset stimuli, the neutral flanker would not attract attention in the small set-size condition. If the attractiveness to attention in the three types of flankers were homogeneous, zero TSED differences (between the flanker and baseline conditions) should be expected with all three of the flanker types. In contrast, while using the offset stimulus, if the neutral flanker were no more attractive to attention, but the congruent and incongruent flankers still attracted attention, this would strongly support the inhomogeneity hypothesis and also reject the possibility that the inhomogeneous attraction might be due to the influence of stimulus–stimulus congruence, because the stimulus–stimulus congruence effect requires that a flanker (irrespective of its identity) be identified in the first place.

Method

Observers An a priori power analysis was conducted with the program G*Power. A repeated measures ANOVA with an effect size of 0.25, α = .05, and $1 - \beta$ = .80 gave a statistical power of 81.98% and a sample size of a minimum of 16 observers. Considering publication bias and uncertainty, 24 students (14 females, 10 males) from Zhejiang University were recruited for Experiment 3. The observers ranged from 18 to 23 years of age (M = 20), with normal or corrected-to-normal vision. None had participated in Experiment 1 or 2.

Design, stimuli, and procedure The design, stimuli, and procedure used in Experiment 3 were identical to those used in Experiment 1, except that offset stimuli were used (Fig. 6). That is, the initial placeholders were block figure-8s rather than short lines, and the letters in the search display emerged from the block figure-8s. On average, 472 trials were completed by each participant in Experiment 3, which lasted about 40 min.

Results

The mean TSEDs for successful searches in the eight stimulus conditions of Experiment 3 are plotted in Fig. 7a, and the mean TSED difference between the baseline and each flanker condition is plotted in Fig. 7b. A two-way (3 flanker types × 2 set sizes) ANOVA conducted on the TSED difference data showed a significant main effect of flanker type, $F(1.61, 37.0)$ = 4.80, p = .02, η_p^2 = .173, but not of set size, $F(1, 23)$ = 2.63, p = .118, η_p^2 = .103. The interaction between flanker type and set size was not significant, $F(2, 46)$ = 0.558, p = .576, η_p^2 = .024. Bonferroni post-hoc tests showed that the TSED difference between the baseline and flanker conditions was significantly larger for the incongruent than for the congruent, p < .01, and neutral, p = .029, flankers. However, the TSED difference for the congruent flanker was not different from that for the neutral flanker, p = 1.0. One-sample t tests further revealed that, in both set-size conditions, the TSED difference was significantly larger than zero only for the incongruent flankers (ps < .05). The mean accuracy rate for trials after the fifth turn in the staircase procedure across all testing conditions was 81.8%, as expected.

Discussion

The crucial finding of Experiment 3 was that even when using offset stimuli, the extents of flanker processing (i.e., the TSED differences between the flanker and baseline conditions) in the three flanker conditions were not identical. For both set-size conditions, processing of the neutral flanker was not significant, but processing of the incongruent flanker was significant. If the different types of flankers attracted attention homogeneously and, with the offset stimuli, the neutral flanker did not attract attention, it would be impossible to explain the significant flanker processing observed in the incongruent flanker condition. Therefore, the findings from Experiment 3 suggested that the three types of flankers indeed attract attention inhomogeneously, although it is not impossible that the significant flanker processing observed in the small set-size conditions of Experiments 1 and 2 might partially be attributed to the stimulus–stimulus congruence effect.
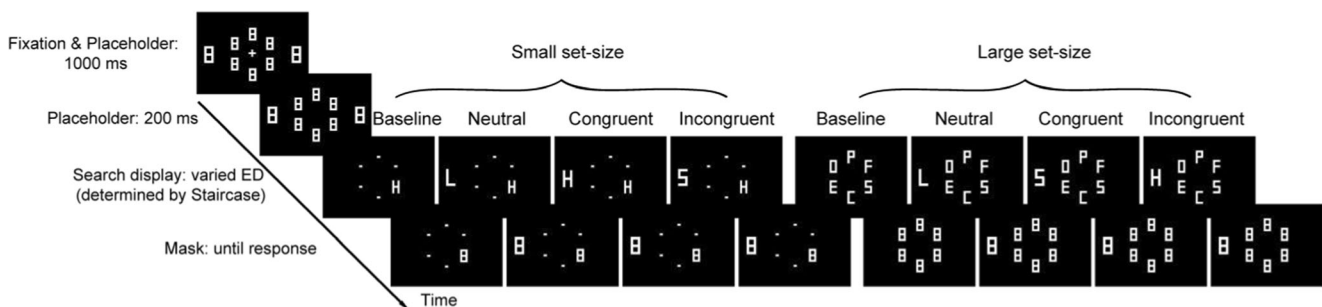
Fig. 6 Paradigm used in Experiment 3. The eight stimulus conditions are plotted in separate columns. Observers judged whether a letter H or S was presented in the central circular array
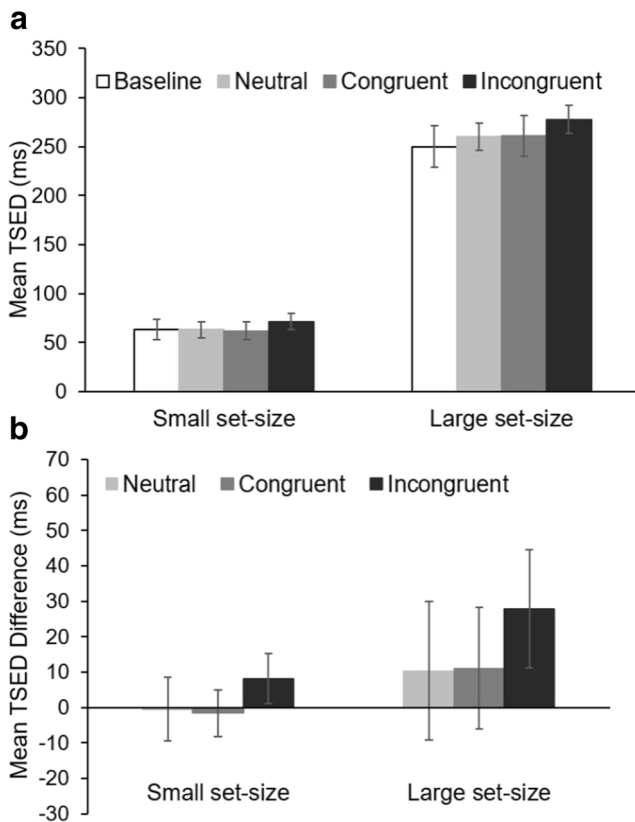
Fig. 7 Results of Experiment 3. (a) Mean threshold stimulus exposure durations (TSEDs) for successfully fulfilling the target identification task in the eight stimulus conditions. (b) Mean TSED difference between the baseline and each flanker condition. Error bars indicate the 95% confidence intervals

General discussion

The main argument we wanted to make is that flanker tasks based on a stimulus–flanker congruency manipulation exaggerate the extent of flanker processing, because the three types of flankers in the congruency manipulation attract attention inhomogeneously. This argument was supported by the results of Experiment 1, that the TSED differences between the baseline and three flanker conditions were different. The fact that, in Experiment 1, the TSED difference for the congruent flanker did not differ from that for the neutral flanker was probably due to an implicit impact of the identity of the congruent flanker on the forced choice response. This assumption was supported by the findings of Experiments 2a and 2b that, when knowing the identity of the congruent flanker could not help the guessing response, the TSED difference became larger for the congruent flanker than for the neutral flanker. Experiment 3 further showed that the inhomogeneous attractiveness of different types of flankers still exists even with offset stimuli, and the present findings cannot solely be explained in terms of the stimulus–stimulus congruence effect.

Although we argue that flanker tasks based on a congruency manipulation are biased measures of distractor

processing in load studies, we do not mean that the findings from previous load studies are untrustworthy. We suggest that future load studies consider applying the present paradigm, because it is not biased when using neutral flankers. Moreover, unlike the traditional flanker paradigm that uses RT as the dependent variable, the present paradigm depends solely on the TSED. Whereas the RT includes the times for stimulus identification, decision making, and executing the response action, the TSED solely depends on the time for stimulus identification. Therefore, the use of TSED makes the present paradigm a very sensitive measure of flanker processing.

It is worth noting that while applying the present new flanker paradigm to examine flanker processing, researchers should use a neutral flanker that does not share the potential identities of the target. If we focus on the TSED data from the neutral flanker conditions of the present study, an interesting finding emerges. In Experiments 1 and 2, processing of the neutral flanker was significant for the small but not for the large set size. At a glance, this is consistent with the perceptual load theory, which proposes that flanker processing is more likely to occur in a low-load task (a small set size is generally considered to be low-load). However, a recent study showed that the available processing time also contributes to perceptual load, so that, with limited processing time, a small set size may also be considered high-load (Li et al., 2018). Because the available processing time in the present study was at the threshold level, both the small and large set-size conditions should be considered high-load. Thus, the present findings actually suggest that, under high load, flanker processing is greater with a small set size. Whereas this finding may not easily be accounted for by the perceptual load theory, it is very consistent with the dilution theory, which suggests that all items in the display attract attention and that the attention allocated to the flanker would be diluted when the display set size is large (Tsal & Benoni, 2010). The finding of Experiment 3 showed that the (neutral) flanker processing disappeared in the small set-size condition when offset stimuli were used, which suggests that the abrupt stimulus onset was what caused the flanker processing, consistent with the activity distribution model of attention (Wright & Ward, 2008).

References

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological Science*, *28*, 1547–1562. https://doi.org/10.1177/09567976177237244

Avital-Cohen, R., & Tsal, Y. (2016). Top-down processes override bottom-up interference in the flanker task. *Psychological Science*, *27*, 651–658.

Broadbent D. E. (1958). *Perception and communication*. London, UK: Pergamon.

Chen, Z., & Cave, K. R. (2016). Zooming in on the cause of the perceptual load effect in the Go/No-go paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1072–1087.

Cohen J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*, 80–90.

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

Eltiti, S., Wallace, D., & Fox, E. (2005). Selective target processing: Perceptual load or distractor salience? *Perception & Psychophysics*, *67*, 876–885. https://doi.org/10.3758/BF03193540

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149. https://doi.org/10.3758/BF03203267

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*, 175–191. https://doi.org/10.3758/BF03193146

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030–1044. https://doi.org/10.1037/0096-1523.18.4.1030

Forster, S., & Lavie, N. (2007). High perceptual load makes everybody equal. *Psychological Science*, *18*, 377–381.

Forster, S., & Lavie, N. (2008). Failures to ignore entirely irrelevant distractors: The role of load. *Journal of Experimental Psychology: Applied*, *14*, 73–83. https://doi.org/10.1037/1076-898X.14.1.73

Johnson, D. N., McGrath, A., & McNeil, C. (2002). Cuing interacts with perceptual load in visual search. *Psychological Science*, *13*, 284–287.

Krueger, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. *Perception & Psychophysics*, *35*, 558–564. https://doi.org/10.3758/BF03205953

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences*, *9*, 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N., & Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychological Science*, *8*, 395–398. https://doi.org/10.1111/j.1467-9280.1997.tb00432.x

Lavie, N., & de Fockert, J. W. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Perception & Psychophysics*, *65*, 202–212. https://doi.org/10.3758/BF03194795

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*, 183–197. https://doi.org/10.3758/BF03213897

Li, Z., Xin, K., Li, W., & Li, Y. (2018). Reconceptualizing perceptual load as a rate problem: The role of time in the allocation of selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 1458–1471.

Lupyan, G. (2008). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition*, *108*, 566–577. https://doi.org/10.1016/j.cognition.2008.03.009

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*, 422–471. https://doi.org/10.1037/0033-295X.110.3.422

Menneer, T., Barrett, D. J. K., Phillips, L., Donnelly, N., & Cave, K. R. (2007). Costs in searching for two targets: Dividing search across target types could improve airport security screening. *Applied Cognitive Psychology*, *21*, 915–932. https://doi.org/10.1002/acp.1305

Menneer, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: Effects of practice and target similarity. *Journal of Experimental Psychology: Applied*, *15*, 125–139.

Scerrati, E., Lugli, L., Nicoletti, R., & Umiltà, C. (2017). Comparing Stroop-like and Simon effects on perceptual features. *Scientific Reports*, *7*, 17815.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Theeuwes, J., Kramer, A. F., & Belopolsky, A. V. (2004). Attentional set interacts with perceptual load in visual search. *Psychonomic Bulletin & Review*, *11*, 697–702. https://doi.org/10.3758/BF03196622

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

Tsal, Y., & Benoni, H. (2010). Diluting the burden of load: Perceptual load effects are simply dilution effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1645–1656. https://doi.org/10.1037/a0018172

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *NeuroImage*, *14*, 1302–1308.

Wilson, D. E., Muroi, M., & MacLeod, C. M. (2011). Dilution, not load, affects distractor processing. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 319–335.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*, 0058. https://doi.org/10.1038/s41562-017-0058

Wright, R. D., & Ward, L. M. (2008). *Orienting of attention*. New York, NY: Oxford University Press.