



Revisiting contextual cueing effects: The role of perceptual processing

Feifei Zhao¹ · Yanju Ren²

Published online: 2 January 2020
© The Psychonomic Society, Inc. 2020

Abstract

Incidental learning of invariant configurations of items can facilitate visual search, which is termed contextual cueing effect. Previous debates about the mechanisms of the contextual cueing effect mainly focused on the roles of attentional guidance and response-related components. Because the time for perceiving displays is very short, few studies have revealed the role of initial perceptual processing. Thus, in the present study, we manipulated the key variable of display contrast to prolong the perceptual processing duration in low-contrast displays. With response-time and eye-movement recording, the study explored the role of initial perceptual processing. In addition, we examined the roles of attentional guidance and response-related components. From three experiments we found: In addition to the role of attentional guidance, the initial perceptual processing could also contribute to a contextual cueing effect in a low-contrast condition; there was only the facilitation from attentional guidance in a high-contrast condition. These results suggest that perceptual processing can also play a role if the perceptual processing time is prolonged, but the main role involves improved attentional guidance.

Keywords Contextual cueing effect · Perceptual processing · Attentional guidance · Visual search · Eye-tracking

Introduction

Objects in a real scene are not arranged randomly but are highly structured. For example, a mouse and a keyboard often appear together; and a pillow is usually on a bed rather than under a bed. Chun and Jiang (1998) found that visual search could be facilitated by learned spatial configurations. In their study, participants were asked to find a T-target among L-distractors. Unknown to the participants, search displays used by half the trials were repeated across blocks, while other search displays were used only once during the whole experiment. There was a benefit with regard to search time in repeated trials compared to novel trials as the experiment progressed, which is termed a *contextual cueing effect*.

Since its inception, a large body of studies have focused on the mechanisms of the contextual cueing effect, specifically, how learned knowledge facilitates visual search. However, the

conclusions have been inconsistent. In general, there are three explanations: attentional guidance hypothesis, faster response-related processing hypothesis, and dual stages hypothesis (both attentional guidance and facilitated response-related processes).

First, the attentional guidance hypothesis supposes that learned configurations can guide attention to targets more efficiently. Studies in favor of this hypothesis are discussed below.

Chun and Jiang (1998) examined the search efficiency of repeated and novel displays (indicated by search slopes and intercepts as a function of varied set size) across three different set sizes. If attentional guidance contributes to a contextual cueing effect, then there should be a greater decrease in search slopes for repeated than for novel displays. However, if other processes contribute to this effect, then the intercepts of search functions should be different in repeated and novel displays. They found a shallower search slope for repeated displays rather than a difference in intercepts, and thus concluded that the contextual cueing effect was driven by attentional guidance.

Harris and Remington (2017) found a reduction in the number of fixations of repeated displays rather than a difference in the time between the last fixation and button press (an indicator of the facilitation from response selection). Furthermore, they used “pop-out” displays, in which

✉ Yanju Ren
renyanju@gmail.com

¹ Laboratory of Emotion and Mental Health, School of Education, Chongqing University of Arts and Sciences, Chongqing, People's Republic of China

² School of Psychology, Shandong Normal University, No. 1 Daxue Road, Changqing District, Jinan 250358, People's Republic of China

attentional guidance was almost maximal. There was still only a reduction in the number of fixations in repeated displays, which suggested a single mechanism, improvements in attentional guidance contributing to the contextual cueing effect. Many other eye-tracking studies also supported this hypothesis (Geringswald, Baumgartner, & Ploomann, 2012; Manelis & Reder, 2012; Peterson & Kramer, 2001; Zang, Jia, Müller, & Shi, 2015).

Previous studies found that N2pc amplitude contralateral to targets was greater for repeated than for novel displays, because N2pc reflects the allocation of attention to the target, thus in favor of the attentional guidance hypothesis (Johnson, Woodman, Braun, & Luck, 2007; Kasper, Grafton, Eckstein, & Giesbrecht, 2015).

Second, the hypothesis of faster response-related processing supposes that contextual cueing is due to improved response selection processes, which need less time between locating targets and committing to responses in repeated than in novel displays.

This assumption was first proposed by Kunar, Flusberg, Horowitz, and Wolfe (2007). Using the same logic as Chun and Jiang (1998), they did not observe a statistical difference on search slopes between repeated and novel displays. However, there was still a stable contextual cueing effect even when the target “popped out” (attentional guidance was regarded as already optimal), although the effect was small. Furthermore, they found that the contextual cueing effect disappeared when the response-related processing was disturbed. Thus, they concluded that contextual cueing effects should be due to faster response selection processes.

Sewell, Colagiuri, and Livesey (2017) contrasted the quantitative predictions of an expedited search account (attentional guidance) and decision threshold account (response-related processing) with a diffusion model. With a typical contextual cueing experiment, they examined how well different versions of the model account for each participant’s response time (RT) data. They found ten of the total 15 participants were best fit by a model that corresponded to a decision threshold account, and only one participant was best fit by a model that corresponded to an expedited search account. The study provided strong support for the latter, and relatively weak support for the former. With the signal detection theory, Schankin, Hagemann, and Schubö (2011) could estimate participants’ sensitivity and response bias. They found that observers showed a more liberal response criterion (beta as the parameter) to repeated displays, but no difference in sensitivity (d' as the parameter); there was a larger amplitude for repeated displays on a late positive activity (reflecting response-related processes) but no difference in the N2pc (reflecting focused attention). These results supported the hypothesis of response-related processing. In addition, studies with incidental learning,

which can be considered as contextual cueing effects in a broad sense, also found that it facilitates resolution of search decisions rather than improving search efficiency (Hout & Goldinger, 2012).

Third, the hypothesis of dual stages suggests that contextual cueing effects benefit from both improved attentional guidance and faster response-related processing. Some evidence supporting this argument is provided below.

With eye-movement recording, Zhao, Liu, Jiao, Zhou, Li, and Sun (2012) divided the total RT of each trial into three segments: the early phase (representing the initial perceptual process, the parameter is initial saccade latency), the middle phase (representing attentional guidance, the value is RT minus the time of early and late phase), and the late phase (representing a response-related process, the parameter is the time between last eye fixation and button press response). They found that the time of the middle phase showed a greater reduction for repeated than for novel displays over epochs, and the difference between repeated and novel displays was more than 100 ms. In addition, the time of late phase also presented a slight decrease on repeated displays compared with novel displays, the difference between them being about 50 ms. These results suggested that attentional guidance contributed most to the contextual cueing effect, but the facilitation of response selection also played a certain role.

Although Kunar et al. (2007) did not observe the role of attentional guidance, Kunar, Flusberg, and Wolfe (2008) suggested that attentional guidance could play a role if it was given sufficient time. We can speculate from both studies that improved attentional guidance as well as faster response-related processing could contribute to contextual cueing effects.

Studies found that both N2pc (reflecting focused attention) and LRP (a component associated with response-related processes) were enhanced in repeated displays, compared with novel displays (Schankin & Schubö, 2009, 2010). This indicated that guidance of attention together with response-related processing account for the entire contextual cueing benefit.

It is well known that we first need to perceive the display when we search for a target in it. So, perceptual processing should be the first phase for a visual search task. One magnetoencephalography (MEG) study found a significant difference between repeated and novel displays in participants’ occipital lobes about 50–100 ms after display onset (Chaumon, Drouet, & Tallon-Baudry, 2008), which suggested that early sensory cortices may have been activated in the contextual cueing effect. In addition, Schlagbauer, Rausch, Zehetleitner, Müller, Geyer, and Notes (2018) found that memory of display context can enhance the representation of the display in the contextual cueing effect. Can the initial perceptual processing then ever play a role in contextual cueing effect? To our knowledge, few studies have examined its role (but see Zhao et al., 2012). In order to ensure that each item had equal

visibility before the first saccade, Zhao et al. (2012) corrected the size of each item according to eccentricity from the fixation point, so that the experiment could be more likely to detect the role of initial perceptual processing. However, they still could not find its role. We think that the failure to find the role of perceptual processing may be due to this phase specifically. The very short period of time for perceptual processing together with the total relatively much longer RTs resulted in them being unable to detect the difference in time between repeated and novel displays easily (of course, we should add, this result may also occur because perceptual processing does not play a role in the contextual cueing effect). Consequently, it is still worth asking whether the initial perceptual processing can ever play a role when the perceptual processing time is prolonged.

The present three experiments were designed to examine this supposition. Because the duration of perceptual processing will be decreased by increasing the contrast of the display (Kojima, & Kawabata, 2012), our experiments test the role of initial perceptual processing in the contextual cueing effect by manipulating the key variable of display contrast. As we know, it takes more time to perceive displays in low than in high contrast. So, if there is a role of perceptual processing in low-contrast displays, then it will also contribute to the benefits of RT for participants searching in repeated displays. Furthermore, because the contextual cueing effect benefits little, if at all, from perceptual processing under typical conditions (high-contrast displays), we might find a stronger contextual cueing effect in low-contrast displays than in high-contrast displays. The direct evidence (also the most powerful) will come from eye-movement data. Initial saccade latency is the time from onset of the display to initiation of the first saccade, and it is influenced by a process of perceptual recognition (Zhao et al., 2012). So, if there is a role of perceptual processing, we would expect shorter initial saccade latencies in repeated than in novel displays. In addition, as previous studies did not reach an agreement on the roles of attentional guidance and response selection, we tested the roles of attentional guidance (by the number of fixations) and response selection (by decision time, the duration between last eye fixation and response) in our experiments. If the effect benefits from improving attentional guidance, there should be fewer fixations required to find targets in repeated displays. Similarly, there should be a shorter time between fixating targets and emitting responses in repeated trials if response-related processing contributes to this effect.

Experiment 1

To examine whether initial perceptual processing has a role, we compared the magnitudes of contextual cueing effect between high- and low-contrast conditions, as well as the initial

saccade latencies between repeated and novel displays. In addition, we examine the roles of attentional guidance and response-related processing.

Method

Participants

Thirty-seven undergraduates (18 men and 19 women, M age = 20.81 years, SD = 1.60 years) took part in the experiment as paid volunteers. Because of unsuitability for G*Power software, the sample size was defined according to the following two factors. First, we referred to some previous highly related studies (Zhao et al., 2012). Second, we increased the sample size appropriately in order to obtain stable results and ensure that enough eye-movement data can be obtained. All participants reported normal or corrected-to-normal vision and had never participated in such a visual search experiment before. They were naïve as to the purpose of this study and gave informed consent prior to their participation.

Apparatus and stimuli

Participants were tested individually in a normally lit room and had their head supported by an eye tracker's chin rest and forehead support about 70 cm from a 19-in. cathode ray tube (iiyama HM903DT B) color monitor (display resolution 1,280 × 1,024 pixels; refresh rate: 85 Hz). E-Prime 2.0 software was used to control event scheduling and collect RTs, run on a PC under the Windows XP operating system. Eye movements were recorded using a Hi-Speed-500 eye-tracking system (Senso Motoric Instruments GmbH, Teltow, Germany) with a sampling rate of 500 Hz and a spatial resolution of 0.01°. An eye movement was classified as a saccade when its velocity reached 40°/s. Minimum fixation duration is 50 ms.

Stimuli were generated by a Matlab Program, similar to those in Chun and Jiang's (1998) Experiment 1, except that we used monochromatic items within each search display. On every trial, either white (RGB: 255, 255, 255; high-contrast trials) or grayish (RGB: 130, 130, 130; low-contrast trials) items were presented against a uniform gray (RGB: 128, 128, 128) background; Michelson contrasts of the stimuli on high- and low-contrast displays were 0.68 and 0.04, respectively (Fig. 1).

Each display contained 11 L-shaped distractors that were rotated 0°, 90°, 180°, or 270° and one T-shaped target rotated 90° to the left or right (balanced across trials). Participants searched for T and pressed one of the two buttons corresponding to the orientation of T. All items were about 1.9° × 1.8° visual angle, and with equal length between horizontal bars and vertical bars. They randomly appeared within an invisible 8 × 6 grid and the center position of each item was slightly

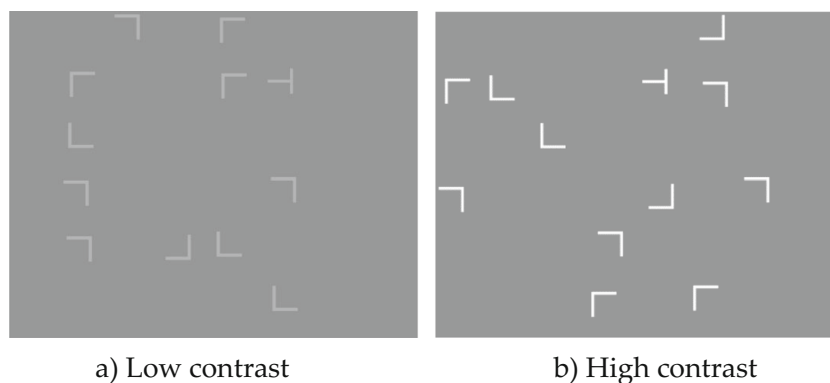


Fig. 1 Examples for the low- (a) and high- (b) contrast displays. In order to show the example of low-contrast displays clearly on a printed version, the color of items in a is grayish (RGB: 160, 160, 160), but in the actual experiment, the color of items in low-contrast displays was a different

gray (RGB: 130, 130, 130). The color of items in b is white (RGB: 255, 255, 255). The background of both kinds of contrast displays was the same gray (RGB: 128, 128, 128)

jittered vertically and horizontally to reduce co-linearity between items. Twelve items were evenly placed into four quadrants of the search displays, with the constraint that items were prevented from appearing in the four corners of displays; in addition, targets could not appear within the central four cells, or in the three cells of each corner.

Procedure and design

Participants were instructed to search for the target (sideways T) and then pressed a button as quickly and as accurately as possible. They were to press the “F” key with their left hand if the target was rotating 90° to the left, and “J” with their right hand if it was rotating 90° to the right. Once they understood the instructions, they completed a practice block of 16 trials. Participants were then calibrated with an eye-tracker’s 9-point calibration, followed by 448 trials (28 blocks of 16 trials) of the formal experiment.

Each trial began with a fixation display. Participants were instructed to stare at the central fixation point for 700 ms. Afterward, the array of stimuli appeared until participants gave a response. The display terminated if no response was given within 9 s. After a brief blank gray screen of 1 s, the next trial began.

The experiment used a within-subject design, and the independent variables were display contrast (high contrast vs. low contrast), display type (repeated vs. novel), and epoch (1–7); the dependent variables were RTs, initial saccade latency (from onset of the display to initiation of the first saccade), number of fixations, and decision time (duration between the last eye fixation and responding). There were 28 blocks of 16 trials each (four repeated and low contrast, four repeated and high contrast, four novel and low contrast, four novel and high contrast). These trials were randomized within blocks. To increase statistical power, blocks were collapsed into seven epochs of four consecutive blocks.

The repeated displays consisted of eight unique search arrays that were repeated across blocks, once per block. The identities of distractors and the locations of all items in these repeated displays were maintained across repetitions, but the identities of targets were randomly chosen to avoid associating a certain response with a certain configuration. The novel displays consisted of eight unique search arrays that were newly generated for each block, with the constraint that the locations of targets were repeated across blocks. The target appeared equally often at each of 16 possible locations throughout the experiment (eight for repeated displays and eight for novel displays). The targets’ locations were evenly distributed across four visual quadrants in both repeated and novel displays.

After the experiment, a surprise recognition task was administered to examine whether observers could explicitly discern repeated from novel displays. This session included 16 displays: eight novel displays (four with high contrast and four with low contrast) were randomly intermixed with eight repeated displays (four displays for each contrast condition) from the earlier visual search task. Participants were instructed to respond to whether the display had appeared in the earlier visual search task without time limit.

Results

In all data analyses in this paper, if the sphericity assumption is violated, a Greenhouse-Geisser correction was applied.

Overall accuracies were very high, over 95%. No difference was observed among display type, display contrast, and epoch in accuracy. Because of a technical error that resulted in a low tracking rate (about 80% or less) of three observers, eye-movement data from the other 34 participants were analyzed further, but the analysis of RTs used data from all 37 participants.

In RTs and eye-tracking analyses, trials were rejected from analysis if their responses were incorrect or the data to be analyzed exceeded three standard deviations of each observer's mean of each condition. This resulted in a loss of 2.61% and less than 2.54% of the data in the RTs analysis and eye-tracking analysis, respectively.

Response times

A 2 (display contrast: high, low) \times 2 (display type: repeated, novel) \times 7 (epoch) repeated-measures analysis of variance (ANOVA) was run on RT data. There were significant main effects of epoch [$F(2.78, 100.13) = 32.36, p < 0.001, \eta_p^2 = 0.47$], display contrast [$F(1, 36) = 156.42, p < 0.001, \eta_p^2 = 0.81$], and display type [$F(1, 36) = 81.40, p < 0.001, \eta_p^2 = 0.69$]. Significant two-way interactions were revealed for epoch \times contrast, $F(3.89, 140.20) = 3.63, p = 0.008, \eta_p^2 = 0.09$, demonstrating a greater downtrend for low contrast as the experiment progressed; epoch \times display type, $F(4.41, 158.87) = 3.06, p = 0.02, \eta_p^2 = 0.08$, indicating the significantly faster responses in repeated than novel conditions as the epoch session progressed; and contrast \times display type, $F(1, 36) = 85.11, p < 0.001, \eta_p^2 = 0.70$, showing different tendencies of RTs on high- and low-contrast displays. The three-way interaction was not significant, $F(6, 216) = 1.53, p = 0.17$. Figure 2a illustrates the mean RTs for each configuration condition as a function of epoch for high- and low-contrast displays.

In order to investigate the magnitude of the contextual cueing effect for different contrast conditions (or in other words, the trends of different display types as the experiment progressed), we performed 2 (display type: repeated, novel) \times 7 (epoch) repeated-measures ANOVAs in high- and low-contrast displays, respectively, although there was no significant three-way interaction of display type, contrast, and epoch. The following eye-movement data were also analyzed in a similar way whether or not there was a significant three-way interaction. For low-contrast trials, there were main effects of epoch [$F(3.67, 132.21) = 18.57, p < 0.001, \eta_p^2 = 0.34$] and display type [$F(1, 36) = 114.95, p < 0.001, \eta_p^2 = 0.76$], and the interaction between epoch and display type was marginally significant, $F(4.36, 156.80) = 2.31, p = 0.06, \eta_p^2 = 0.06$; RTs were faster for repeated displays ($M = 1,312$ ms) than for novel displays ($M = 1,490$ ms), indicating the presence of typical contextual cueing effect. A marginally significant interaction is common due to contextual cueing emerging already at the first epoch (e.g., Chun & Jiang, 1998; Harris & Remington, 2017). For high-contrast trials, there was a significant main effect of epoch [$F(2.08, 74.79) = 40.42, p < 0.001, \eta_p^2 = 0.53$] and a significant interaction between epoch and display type [$F(4.27, 153.81) = 2.68, p = 0.03, \eta_p^2 = 0.07$]; the main effect of display type was not significant, $F(1, 36) = 1.31, p = 0.26$. Further analysis of the interaction only

revealed significantly longer RTs of repeated trials in epoch 1, so we did not observe a contextual cueing effect on high-contrast displays.

Recognition test

Mean accuracy in the explicit recognition task was 54.42%; this differed significantly from a chance guessing level of 50%; two-tailed $t(36) = 2.56, p = 0.015$, Cohen $d = 0.85$, suggesting that a contextual cueing effect may not be implicit. However, when we compared the mean accuracy of the low- and high-contrast conditions separately with 50%, the former revealed only a marginally significant difference $t(36) = 1.83, p = 0.075$, Cohen $d = 0.61$, two-tailed, while the difference with the latter was not significant, $t(36) = 1.36, p = 0.18$, Cohen $d = 0.45$. In addition, the correlation between a contextual cueing effect and memory score was not significant, $r = -0.05, p = 0.77$, suggesting that a contextual cueing effect is implicit to some extent. Considering that there were controversial opinions on the explicit or implicit nature of contextual memory (Vadillo, Konstantinidis, & Shanks 2016), we do not draw any strong conclusions on this issue.

Eye movements

We analyzed three eye-tracking parameters: initial saccade latency, the time from onset of the display to initiation of the first saccade, our proxy for perceptual processing; the number of fixations corresponding to the attentional guidance; and decision time, duration between the last eye fixation and responding, our proxy for response-related processing (Zhao et al., 2012).

Initial saccade latency A 2 (display contrast: high, low) \times 2 (display type: old, new) \times 7 (epoch) repeated-measures ANOVA revealed a significant main effect of display contrast, $F(1, 33) = 303.77, p < 0.001, \eta_p^2 = 0.90$, the interaction between display contrast and display type [$F(1, 33) = 26.54, p < 0.001, \eta_p^2 = 0.45$] was also significant, indicating different tendencies of initial saccade latencies on high- and low-contrast displays. In addition, the interaction between epoch and display contrast was significant, $F(6, 198) = 5.61, p < 0.001, \eta_p^2 = 0.15$, suggesting that the tendencies of initial saccade latencies were different on high- and low-contrast displays as the experiment progressed. None of the other main effects or interactions were significant, $ps > 0.29$. We also conducted 2 (display type) \times 7 (epoch) repeated-measures ANOVAs on high- and low-contrast displays, respectively (Fig. 2b). For low-contrast trials, a significant main effect of display type was observed [$F(1, 33) = 8.89, p = 0.005, \eta_p^2 = 0.21$], suggesting that the initial saccade latencies were shorter in repeated than in novel displays, but there was no main effect of epoch or the interaction of both factors. The lack of

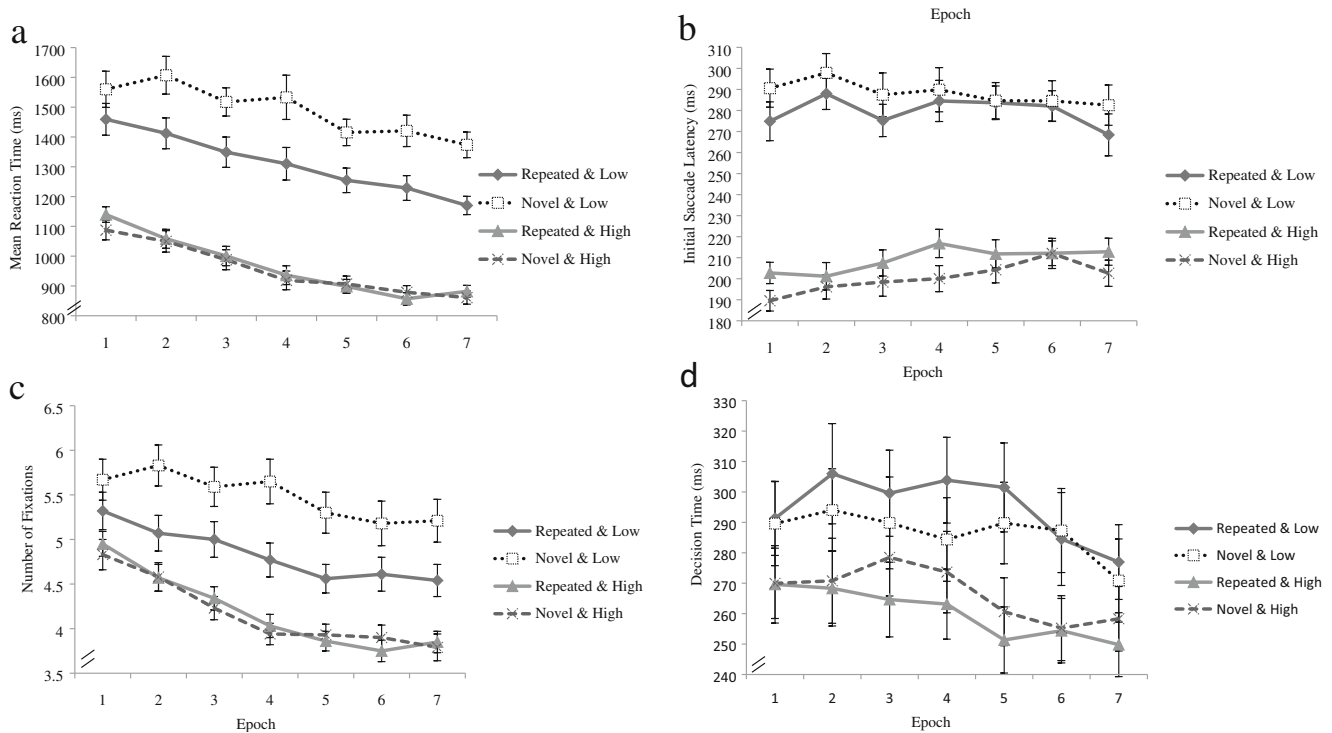


Fig. 2 Results of Experiment 1. **(a)** Mean reaction times for searching targets. **(b)** Initial saccade latency. **(c)** Number of fixations before response. **(d)** Time between the start of the last fixation and behavioral response, as a function of epoch in four variable combinations, respectively. For the Repeated & Low condition, a repeated display was

presented with low contrast; for the Novel & Low condition, a novel display was presented with low contrast; for the Repeated & High condition, a repeated display was presented with high contrast; for the Novel & High condition, a novel display was presented with high contrast. Error bars represent the within-subject standard error

interaction is puzzling, and we explore the possible reason in the *Discussion* section. For high-contrast displays, there were significant main effects of epoch [$F(3.98, 131.38) = 4.04, p = 0.004, \eta_p^2 = 0.11$] and display type [$F(1, 33) = 33.00, p < 0.001, \eta_p^2 = 0.5$], while their interaction was not significant ($p = 0.27$), suggesting that the initial saccade latencies of repeated displays were longer than novel displays, and the initial saccade latency presented tenuous increasing as the experiment progressed.

Number of fixations We first performed a 2 (display contrast: high, low) \times 2 (display type: old, new) \times 7 (epoch) repeated-measures ANOVA. This revealed significant main effects of all three factors: epoch, $F(3.76, 124.00) = 36.64, p < 0.001, \eta_p^2 = 0.53$; display contrast, $F(1, 33) = 63.58, p < 0.001, \eta_p^2 = 0.66$; display type, $F(1, 33) = 65.32, p < 0.001, \eta_p^2 = 0.66$. The interaction between epoch and display contrast was significant, $F(6, 198) = 7.34, p < 0.001, \eta_p^2 = 0.18$, as was the interaction between display contrast and display type, $F(1, 33) = 45.81, p < 0.001, \eta_p^2 = 0.58$. The interaction between epoch and display type was marginally significant, $F(3.93, 129.61) = 2.33, p = 0.06, \eta_p^2 = 0.07$, and the three-way interaction was also marginally significant, $F(6, 198) = 2.03, p = 0.06, \eta_p^2 = 0.06$. We then conducted 2 (display type) \times 7 (epoch) ANOVAs on high- and low-contrast displays,

respectively (Fig. 2c). For low-contrast trials, this revealed a significant main effect of epoch, $F(4.36, 143.72) = 12.76, p < 0.001, \eta_p^2 = 0.28$, a significant main effect of display type, $F(1, 33) = 85.86, p < 0.001, \eta_p^2 = 0.72$, and a significant interaction between them, $F(4.30, 141.95) = 2.52, p = 0.04, \eta_p^2 = 0.07$. The results suggested that fewer fixations were gradually required to find targets on repeated than on novel displays as the experiment progressed. For high-contrast trials, there was only a significant main effect of epoch, $F(3.83, 126.45) = 53.15, p < 0.001, \eta_p^2 = 0.62$, neither the main effect of display type [$F(1, 33) = 0.13, p = 0.72$] nor the interaction between epoch and display type [$F(4.18, 137.76) = 1.61, p = 0.17$] was significant.

Decision time Similarly, a 2 (display contrast: high, low) \times 2 (display type: old, new) \times 7 (epoch) repeated-measures ANOVA was first conducted. This revealed significant main effects of epoch [$F(3.77, 124.39) = 2.86, p = 0.03, \eta_p^2 = 0.08$] and display contrast [$F(1, 33) = 28.86, p < 0.001, \eta_p^2 = 0.47$]; the interaction between display contrast and display type was also significant [$F(1, 33) = 10.78, p = 0.002, \eta_p^2 = 0.25$]. Neither the main effect of display type nor the other interactions was significant, $ps > 0.35$. We also conducted 2 (display type) \times 7 (epoch) repeated-measures ANOVAs on high- and low-contrast displays, respectively (Fig. 2d). For low-contrast

displays, there were significant main effects of epoch [$F(6, 198) = 2.38, p = 0.03, \eta_p^2 = 0.07$] and display type [$F(1, 33) = 6.38, p = 0.02, \eta_p^2 = 0.16$]. The interaction between them was not significant, $F(6, 198) = 0.66, p = 0.68$. These results suggest that decision time was longer on repeated than on novel displays. For high-contrast displays, the main effects of epoch [$F(4.16, 137.15) = 2.10, p = 0.08, \eta_p^2 = 0.06$] and display type [$F(1, 33) = 3.62, p = 0.07, \eta_p^2 = 0.10$] were all marginally significant. Their interaction was also not significant, $F(4.29, 141.58) = 0.57, p = 0.70$. The decision time of repeated displays was shorter than novel displays.

Discussion

The results here clearly show the facilitation of initial perceptual processing as a source of contextual cueing effect in low-contrast displays. The magnitude of contextual cueing in low contrast was larger than in high contrast: We observed the effect under a low-contrast condition, but not under a high-contrast condition. Furthermore, the eye-tracking data revealed that repeated displays shorten the initial saccade latencies under a low-contrast condition compared with novel displays. All the evidence converges on the conclusion that perceptual processing also contributes to a contextual cueing effect if its time is prolonged.

In agreement with most previous studies (Chun & Jiang, 1998; Harris & Remington, 2017; Johnson et al., 2007; Zang et al., 2015), this experiment also suggested that improved attentional guidance contributed to a contextual cueing effect (here we refer to low-contrast displays only, since there was no contextual cueing effect on high-contrast displays). Repeated displays reduced the number of fixations required to find targets. We didn't find a benefit from response selection on low-contrast displays, in fact there was a slight cost. Repeated displays showed a slightly longer time between the last eye fixation and key-press response in the low-contrast condition. We discuss this phenomenon in detail in the *General discussion*.

In sum, Experiment 1 suggested that both initial perceptual processing and attentional guidance contributed to the contextual cueing effect, at least in the low-contrast condition.

One puzzling aspect of the results is the lack of interaction between epoch and display type in low-contrast trials for initial saccade latencies. As a phase of the search task, initial perceptual processing should have shown a significant interaction since contextual cueing only emerged after statistical learning of the displays had been established in early epochs. However, previous studies have suggested that contextual cueing sometimes can emerge within the first epoch of the experiment (Chun & Jiang, 1998; Geyer, Zehetleitner, & Müller, 2010; Harris & Remington, 2017; Kunar et al., 2007, 2008; Zhao et al.,

2012). In the same way, we speculate that it is a common result due to the role of initial perceptual processing emerging within the first epoch. As with Chun and Jiang (1998), we compared the difference between repeated and novel configurations of the first few blocks on the initial saccade latencies (Fig. 3). We found that the benefit of the initial saccade latency was not significant until block 5 [two-tailed $t(33) = 2.09, p = 0.04, \text{Cohen } d = 0.73$], in fact block 4 had been marginally significant [two-tailed $t(33) = 1.88, p = 0.07, \text{Cohen } d = 0.65$], so the benefit did not exist from the start. In addition, the duration of the phase of initial perceptual processing per se is very short, and this meant any small factors could influence the statistical results, such as the small number of trials within each block (16 trials here, so only eight configurations need to be remembered), which makes learning earlier and easier.

The loss of contextual cueing under a high-contrast condition was surprising. Eye-movement data showed that the benefit of repeated displays from response selection was nullified by the cost of initial perceptual processing; there was no benefit from attentional guidance (Fig. 2). The results may be accounted for by interference from the low-contrast trials. In Experiment 1, low- and high-contrast displays intermixed within blocks and appeared randomly, but they were very different with regard to difficulty of search tasks. As we can see, the average RT was 1,383 ms for low-contrast trials, whereas it was 965 ms for high-contrast trials. Consequently, we assumed that participants focused most of their cognitive resources on the more difficult displays (low-contrast displays). Although contextual cueing, as an incidental learning or statistical learning phenomenon, can occur with minimal attentional resources, the learning effect is usually weaker than when attention is fully available (Musz, Weber, & Thompson-Schill, 2015; Pollmann, 2019; Thiessen, Kronstein, & Hufnagle, 2013).

In addition, the target search task of this experiment, especially for the high-contrast trials, is easier overall than that in most of the previous studies. First, the stimuli were similar to those in Chun's Experiment 1 (Chun & Jiang, 1998), except that we used monochromatic items, but the size of our stimuli was bigger in the visual angle. Second, there were only 16 trials in each block (eight trials in the high-contrast condition). As a whole, this resulted in a relatively fast search time. We can speculate from previous studies that the magnitude of the contextual cueing effect decreases as the difficulty of the search task decreases (Geyer, Zehetleitner, & Müller, 2010; Harris & Remington, 2017; Kunar, Flusberg, & Wolfe, 2006). Indeed, Harris and Remington (2017) did not observe evidence of contextual cueing following valid spatial cues in their first two experiments. So, we supposed that weaker contextual cueing benefits per se together with the interference of low-contrast trials resulted in a null significant effect for high-contrast trials. Experiment 2 tested this possibility.

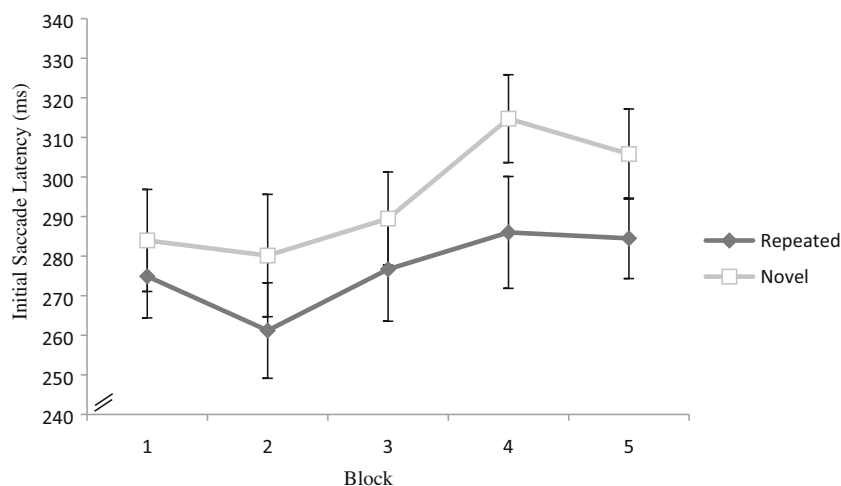


Fig. 3 The initial saccade latency in low-contrast trials as a function of the first five blocks. For the Repeated condition a repeated display was used, while for the Novel condition a novel display was presented. Error bars represent the within-subject standard error

Experiment 2

As noted above, the loss of contextual cueing for high-contrast trials in Experiment 1 is probably due to interference from low-contrast trials. To test this possibility and examine the reliability of the role of initial perceptual processing in contextual cueing effect, we ran this experiment, identical to Experiment 1 except that high-contrast displays were presented separately from low-contrast displays. Participants first completed one type of contrast display and then finished the other type of contrast display. Thus, if it is the interference from low-contrast trials that resulted in the loss of contextual cueing in high-contrast trials, we should observe the contextual cueing effect in this experiment. Critically, we can verify the account that contextual cueing also benefits from expedited perceptual processing of repeated displays.

Method

Participants

Thirty-one new undergraduates (10 men and 21 women, M age = 20.16 years, SD = 1.16 years) were paid to participate in Experiment 2. We reduced the sample size slightly relative to Experiment 1 because participants could complete Experimental 1 easily. All participants had normal or corrected-to-normal vision and had never participated in such a visual search experiment before. They were naïve as to the purpose of this study and gave informed consent prior to their participation.

Stimuli and procedure

The equipment and stimuli were identical to the ones used in Experiment 1 except that the visual search task

was completed in two sessions. One session contained all high-contrast displays, while the other session contained only low-contrast displays. So, each session consisted of 28 blocks of eight trials each (four repeated displays, four novel displays). The order of the two sessions was shuffled so that half the participants finished the high-contrast trials first and then finished the session of low-contrast displays, and this order was reversed for the other half of participants. In practice trials, participants were presented with only one contrast level, which was the same as the first session displays. Following the first session, participants were given a break of at least 5 min but were allowed to wait as long as they wished before beginning the second session. Consistent with Experiment 1, after the visual search tasks of two sessions, participants performed a recognition test that contained high- and low-contrast displays from the two sessions.

Results

For visual search task, the overall accuracy was very high, above 99%. A 2 (display contrast) \times 2 (display type) \times 7 (epoch) repeated-measures ANOVA on accuracy only revealed a main effect of display type, $F(1, 30) = 12.61$, $p = 0.001$, $\eta_p^2 = 0.30$, demonstrating that accuracy was higher for repeated than for novel displays.

Trials with incorrect responses were excluded from the analysis, which resulted in a loss of 0.78% of all trials. For each experimental condition, RTs and eye-tracking data outside the range of ± 3 standard deviations from means were also discarded as “outliers” (1.63% and less than 1.48%, respectively).

Reaction times

A 2 (display contrast: high, low) × 2 (display type: repeated, novel) × 7 (epoch) repeated-measures ANOVA on the RT data revealed a significant main effect of display contrast, $F(1, 30) = 125.86, p < 0.001, \eta_p^2 = 0.81$, a significant main effect of display type, $F(1, 30) = 72.16, p < 0.001, \eta_p^2 = 0.71$, and a significant main effect of epoch, $F(6, 180) = 43.53, p < 0.001, \eta_p^2 = 0.59$. The interaction between epoch and display type was significant, $F(6, 180) = 3.67, p = 0.002, \eta_p^2 = 0.11$, as was the interaction between display contrast and display type, $F(1, 30) = 32.98, p < 0.001, \eta_p^2 = 0.52$. The interaction between epoch and display contrast was marginally significant, $F(3.71, 111.14) = 2.15, p = 0.09, \eta_p^2 = 0.07$. There was no significant three-way interaction between display contrast, display type, and epoch, $F(6, 180) = 1.71, p = 0.12$. Figure 4a illustrates the mean RT values for each configuration condition as a function of epoch for high and low contrast.

Analogous to Experiment 1, to compare the contextual cueing effect for different display contrasts, we conducted separate ANOVAs (display type by epoch) on high- and low-contrast trials. For low-contrast trials, there was a significant main effect of epoch, $F(6, 180) = 21.91, p < 0.001, \eta_p^2 = 0.42$, and a significant main effect of display type, $F(1, 30) = 69.71, p < 0.001, \eta_p^2 = 0.70$. The interaction between epoch and display type was also significant, $F(4.67, 140.01) = 2.80,$

$p = 0.022, \eta_p^2 = 0.09$, demonstrating that participants were able to respond significantly faster in repeated ($M = 1,230$ ms) than in novel displays ($M = 1,349$ ms) as the experiment progressed. For high-contrast trials, we found a significant main effect of epoch, $F(3.26, 97.78) = 18.81, p < 0.001, \eta_p^2 = 0.39$, and interaction of epoch and display type, $F(6, 180) = 2.70, p = 0.016, \eta_p^2 = 0.08$. There was no significant main effect of display type, $F(1, 30) = 1.65, p = 0.21$. This demonstrated that the contextual cueing effect still do not generate steadily in high-contrast displays in spite of a trend for contextual cueing in epochs 4, 5, and 7.

Recognition test

The results of the recognition test replicated those obtained in Experiment 1, revealing a significant difference between mean accuracy (54.64%) and the guessing level of 50%, $t(30) = 3.02, p = 0.005, \text{Cohen } d = 1.10$, suggesting that the contextual cueing effect may not be implicit. While mean accuracy in low-contrast trials was marginally significantly different with guessing level $t(30) = 1.79, p = 0.08, \text{Cohen } d = 0.65$, for high-contrast trials the difference was not significant, $t(30) = 1.65, p = 0.11, \text{Cohen } d = 0.60$. The correlation between contextual cueing effect and memory score was not significant, $r = 0.22, p = 0.24$, suggesting that a contextual cueing effect is implicit.

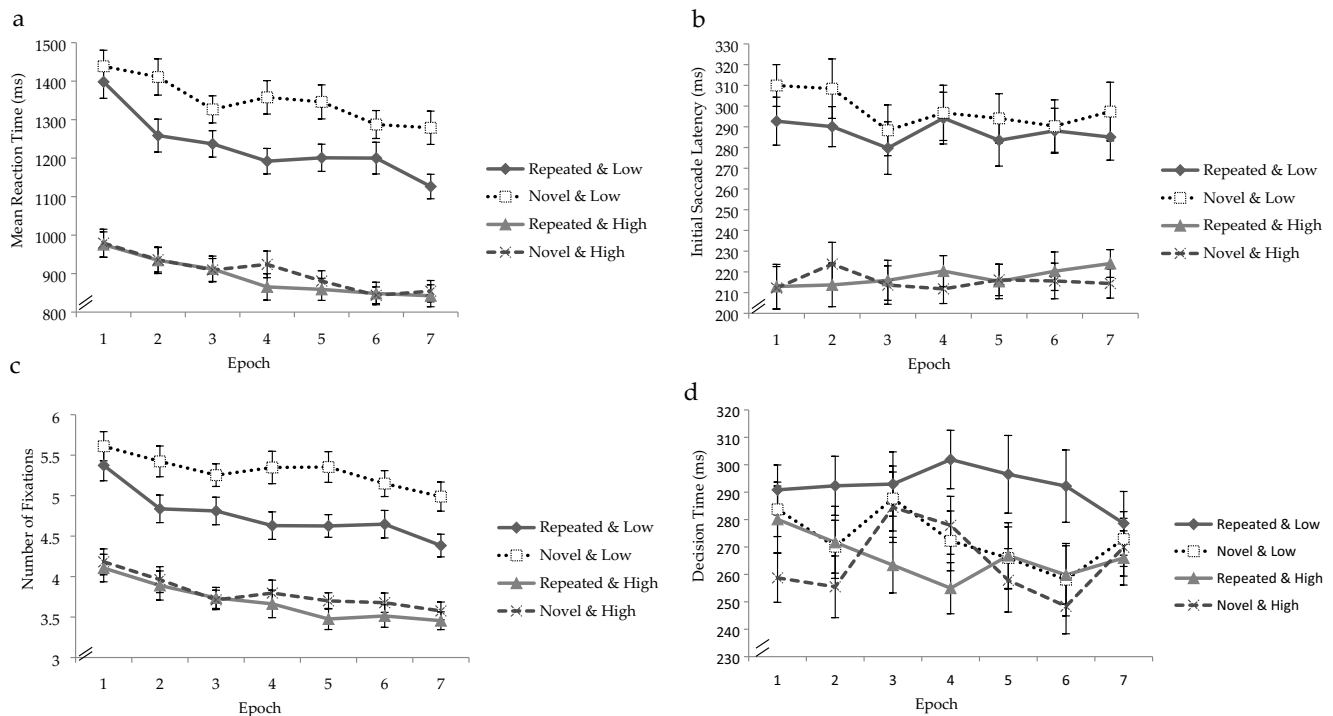


Fig. 4 Results of Experiment 2. (a) Mean reaction times for searching targets. (b) Initial saccade latency. (c) Number of fixations before response. (d) Time between the start of the last fixation and behavioral response, as a function of epoch in four variable combinations, respectively. For the Repeated & Low condition, a repeated display was

presented with low contrast; for the Novel & Low condition, a novel display was presented with low contrast; for the Repeated & High condition, a repeated display was presented with high contrast; for the Novel & High condition, a novel display was presented with high contrast. Error bars represent the within-subject standard error

Eye movements

Because of a technical error, three participants' data were not recorded. Eye-tracking data of the other 28 participants were analyzed further.

Initial saccade latency A 2 (display type) \times 2 (display contrast) \times 7 (epoch) repeated-measures ANOVA revealed a significant main effect of display contrast, $F(1, 27) = 148.12, p < 0.001, \eta_p^2 = 0.85$, and a significant main effect of display type, $F(1, 27) = 8.25, p = 0.008, \eta_p^2 = 0.23$. The interaction between display contrast and display type was also significant, $F(1, 27) = 17.41, p < 0.001, \eta_p^2 = 0.39$. None of the other main effects or interactions were significant, $ps > 0.18$.

We conducted separate ANOVAs (display type by epoch) on high and low contrast (Fig. 4b). There was a significant main effect of display type for low-contrast trials [$F(1, 27) = 19.85, p < 0.001, \eta_p^2 = 0.42$], reflecting shorter initial saccade latencies in repeated displays, without a significant main effect of epoch ($p > 0.20$) or interaction between them ($p > 0.66$). As with Experiment 1, we compared the difference of initial saccade latencies between repeated and novel configurations within the first few blocks. This revealed that participants did not show any contextual cueing benefit in block 1 [$t(27) = 0.96, p = 0.35, \text{Cohen } d = 0.37$], but did exhibit it to some extent in block 2 [$t(27) = 1.94, p = 0.06, \text{Cohen } d = 0.75$]. For high-contrast trials, there was no significant main effect or interaction: epoch, $F < 1, p > 0.77$; display type, $F = 2.59, p = 0.12$; epoch by display type, $F = 1.76, p = 0.11$. Overall, the results collectively suggested again that shorter initial saccade latency contributes to RT benefits in repeated displays.

Number of fixations

A 2 (display contrast) \times 2 (display type) \times 7 (epoch) repeated-measures ANOVA showed a significant main effect of epoch, $F(3.36, 90.76) = 23.22, p < 0.001, \eta_p^2 = 0.46$, a significant main effect of display contrast, $F(1, 27) = 94.81, p < 0.001, \eta_p^2 = 0.78$, and a significant main effect of display type, $F(1, 27) = 80.11, p < 0.001, \eta_p^2 = 0.75$. The interaction between display contrast and display type was significant, $F(1, 27) = 31.18, p < 0.001, \eta_p^2 = 0.54$, as well as the interaction between epoch and display type, $F(6, 162) = 2.20, p = 0.05, \eta_p^2 = 0.08$. None of the other interactions was significant, all $p > 0.45$.

We performed 2 (display type: repeated, novel) \times 7 (epoch) repeated-measures ANOVAs in high- and low-contrast trials, respectively (Fig. 4c). The ANOVA of low-contrast trials yielded a significant main effect of epoch, $F(3.78, 102.03) = 11.23, p < 0.001, \eta_p^2 = 0.29$, and a significant main effect of display type, $F(1, 27) = 78, p < 0.001, \eta_p^2 = 0.74$, but the interaction of them was only marginally significant, $F(6, 162) = 1.80, p = 0.10, \eta_p^2 = 0.06$. Analogous to the initial saccade latency, we consider this lack of interaction a common finding

due to the role of attentional guidance emerging within the first epoch, or other factors such as the tiny difference per se or the small numbers of trials within each block. The results produced by ANOVA for high-contrast trials were the same as for low-contrast trials. There were significant main effects of epoch [$F(3.10, 83.79) = 13.57, p < 0.001, \eta_p^2 = 0.33$] and display type [$F(1, 27) = 5.9, p = 0.022, \eta_p^2 = 0.18$], but no interaction between them, $F(6, 162) = 1.22, p = 0.30$. These revealed that fewer fixations were required to find the targets for repeated displays in both high and low contrast.

Decision time

A 2 (display contrast) \times 2 (display type) \times 7 (epoch) repeated-measures ANOVA on the time between the last eye fixation and responding revealed a significant main effect of epoch, $F(6, 162) = 2.31, p = 0.04, \eta_p^2 = 0.08$, a significant main effect of display type, $F(1, 27) = 10.45, p = 0.003, \eta_p^2 = 0.28$, and a marginally significant main effect of display contrast, $F(1, 27) = 3.31, p = 0.08, \eta_p^2 = 0.11$. The interaction between epoch and display type was significant, $F(6, 162) = 3.17, p = 0.006, \eta_p^2 = 0.11$, as was the interaction between display type and display contrast, $F(1, 27) = 15.07, p = 0.001, \eta_p^2 = 0.36$. There was no significant interaction between epoch and display contrast ($p > 0.95$). The three-way interaction between epoch, display type, and display contrast was also significant, $F(6, 162) = 2.93, p = 0.01, \eta_p^2 = 0.10$.

We followed up the three-way interaction with separate ANOVAs (display type by epoch) on high and low contrast (Fig. 4d). The ANOVA of low-contrast trials only revealed a significant main effect of display type, $F(1, 27) = 22.36, p < 0.001, \eta_p^2 = 0.45$, as in Experiment 1, suggesting that decision time was significantly longer on repeated than on novel displays. Both the main effect of epoch ($p = 0.40$) and the interaction between epoch and display type ($p = 0.14$) were not significant. For high-contrast trials, there was only a significant interaction between epoch and display type, $F(6, 162) = 5.23, p < 0.001, \eta_p^2 = 0.16$, revealing a longer decision time on repeated displays in the first two epochs, and a shorter decision time on repeated displays in epoch 4 and epoch 5. The main effects of epoch ($p = 0.29$) and display type ($p = 0.71$) were not significant.

Discussion

Experiment 2 replicated the main results of Experiment 1, finding a steadily contextual cueing effect only in low-contrast displays and benefits on the initial saccade latency as well as the number of fixations in repeated displays. These again demonstrate that the initial perceptual processing could be a driving force of the contextual cueing effect, provided that participants take sufficient time to perceive the displays; in addition, attentional guidance also plays a major role.

The improved attentional guidance found is consistent with a large number of previous studies (Chun & Jiang, 1998; Harris & Remington, 2017; Zhao et al., 2012).

One of the results that bears commenting on is the decision time, our proxy for the duration of response-related processes. Although there was no significant difference on this indicator between repeated and novel displays with high contrast, which is a common finding (e.g., Chun & Jiang, 1998; Harris & Remington, 2017), decision times were longer on repeated than on novel displays with low contrast. This result is consistent with Experiment 1, which also showed a not very reliable cost. We will try to explain these divergent results in detail in *General discussion*.

Another confusing and also most important issue is that there is still no steadily contextual cueing effect in high-contrast displays. Does the interference from low-contrast trials still persist? We first examined this possibility. If it was the case, the contextual cueing effect of high-contrast displays should larger for those participants who performed them first than those who performed them second. The magnitude of a contextual cueing effect was calculated according to the data of the last four epochs. Our analysis found no significant difference between them (the former is 16 ms, the latter is 27 ms), $t(29) = 0.53$, $p = 0.60$, Cohen $d = 0.20$. This suggested that the manipulation of Experiment 2 eliminated the interference of low-contrast trials successfully, and there are probably other reasons for the weak contextual cueing effect under the high-contrast condition.

We also compared the magnitude of the contextual cueing effect in the first session ($M = 78$ ms) with that in the second session ($M = 82$ ms); the results revealed no significant difference, $t(30) = 0.13$, $p = 0.90$, Cohen $d = 0.05$. Lastly, we compared the contextual cueing effect of low-contrast trials for participants who performed them first ($M = 136$ ms) with those who performed them second ($M = 140$ ms), and the difference was still not significant, $t(29) = 0.11$, $p = 0.91$, Cohen $d = 0.04$. Although no T-test of the three pairs revealed a significant difference, we found a trend that contextual cueing effect of the second session was slightly bigger than the first session, whether in high or low contrast.

Higuchi and Saiki (2017) found that the contextual cueing effect occurred earlier and was bigger to some extent when participants restricted their eye movements compared to when eye movements were allowed. This suggests that the strategy of searching used by participants may have an effect on the contextual cueing effect. In addition, studies suggested that both stress and negative emotion could impair the contextual cueing effect (Kunar, Watson, Cole, & Cox, 2014b; Meyer, Quaedflieg, Bisby, & Smeets, 2019). Thus, we speculated that participants may be more focused (or active) during the first session search task because of emotions such as expectations compared with the second session task. In addition, with the eye-tracker chin rest, participants' heads were fixed, and this

may have resulted in mental stress. The specific participants' mental states could decrease the magnitude of the contextual cueing effect and result in the loss of the already small contextual cueing with high contrast.

Although the above-mentioned T-test analysis suggested that the interference from low-contrast trials was eliminated. But the null significant difference may have resulted from reductive power by splitting the data. So, interference from low-contrast trials could still exist.

So, it was necessary to conduct experiment 3, which excludes the possible interference of low-contrast trials and the factor of variation in participants' mental state.

Experiment 3

As noted above, in order to rule out possible interference, Experiment 3 examined the contextual cueing effect only with high contrast; in order to control the factor of participants' mental state, this experiment tested only at the level of RTs. By this means, we wanted to make sure that the contextual cueing effect can occur in our experiment as in most of previous studies.

Method

Participants

Thirty-three new undergraduates (11 men and 22 women, M age = 20.33 years, $SD = 1.24$ years) were paid to participate in Experiment 3. All had normal or corrected-to-normal vision and had never participated in such a visual search experiment before. They were naïve as to the purpose of this study and gave informed consent prior to their participation.

Stimuli and procedure

The equipment and stimuli were identical to those in Experiment 1 except that we no longer used the eye tracker (heads no longer were fixed), and there were only high-contrast trials. So, the experiment consisted of 28 blocks of eight trials each (four repeated displays, four novel displays). Similarly, after the visual search tasks, participants performed a recognition test.

Results

The overall accuracy of the visual search task was very high, above 95%. No difference was observed between display type and epoch in accuracy.

Trials with incorrect responses were excluded from the analysis, which resulted in a loss of 1.18% of all trials. For each experimental condition, RTs outside the range of ± 3

standard deviations from the means were also discarded as “outliers” (1.42%).

Reaction times

A 2 (display type: repeated, novel) \times 7 (epoch) repeated-measures ANOVA on the RT data revealed a significant main effect of epoch [$F(4.23, 135.34) = 57.50, p < 0.001, \eta_p^2 = 0.64$] and a marginally significant main effect of display type, $F(1, 32) = 3.25, p = 0.08, \eta_p^2 = 0.09$. The interaction between epoch and display type was also significant, $F(6, 192) = 2.76, p = 0.01, \eta_p^2 = 0.08$, suggesting that RTs on repeated displays were gradually shorter than on novel displays as the experiment progressed. Although the results of this experiment were similar to results obtained in Experiment 2 statistically (the only difference between them was that the main effect of display type changed from not significant to marginally significant), the trends of the data now were more consistent with most previous studies (Chun & Jiang, 1998) (see Fig. 5).

Recognition test

The results revealed a significant difference between mean accuracy (60.99%) and the guessing level of 50%, $t(32) = 4.66, p < 0.001$, Cohen $d = 1.65$, suggesting that a contextual cueing effect may be explicit. But the correlation between contextual cueing effect and memory score was not significant, $r = 0.22, p = 0.21$, suggesting that memory of displays may be implicit.

Discussion

The results here clearly revealed a contextual cueing effect with high contrast, although it was still very weak, about 31 ms according RTs of the last four epochs. Thus, it demonstrates our presumption that the unstable contextual cueing effect of high contrast in Experiment 2 was due to interference from low-contrast trials and/or participants’ specific state.

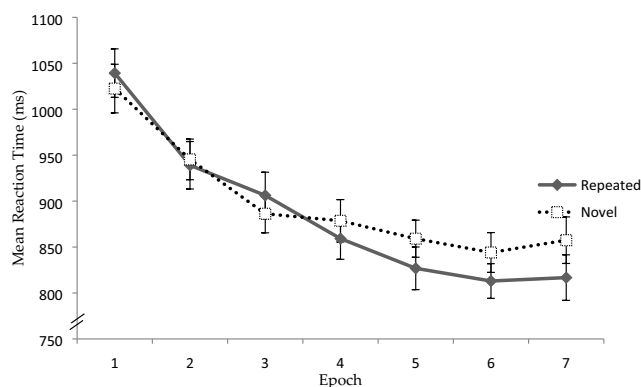


Fig. 5 Results of Experiment 3. Mean reaction times for searching targets as a function of epoch in repeated and novel displays respectively. Error bars represent the within-subject standard error

It is worth noting that the contextual cueing effect of this experiment is really small comparing with previous classic studies, such as that of Chun and Jiang (1998), in which the magnitude of contextual cueing in Experiment 1 was 71 ms. As we noted in the discussion of Experiment 1, the search task was easier than Experiment 1 of Chun and Jiang (1998), so the weak contextual cueing is reasonable.

Another point to note is that the two factors that we presumed resulted in unstable contextual cueing of Experiment 2 on high contrast did not affect (or if at all, only a very small effect) the data of low contrast. In our opinion, it is the very weak contextual cueing in high contrast that caused the two factors that could have an effect.

General discussion

In order to examine the mechanisms of contextual cueing effect, particularly the role of initial perceptual processing, this study lengthened the time for perceptual processing by manipulating the contrast of search displays. With an eye-tracking technique, we found that initial perceptual processing could also contribute to a contextual cueing effect, provided that the perceptual processing time is prolonged. The other driving force of contextual cueing effect was attentional guidance. However, we did not observe the role of response-related processing. In low contrast, participants need shorter initial saccade latency and fewer fixations to locate targets for repeated than for novel displays. Namely, both initial perceptual processing and attentional guidance contribute to the contextual cueing effect. With high contrast the contextual cueing effect benefited solely from reduction of the number of fixations, which means that only attentional guidance facilitates the visual search.

This study extended the mechanisms of contextual cueing effect further and clearly showed that perceptual processing could also contribute to the contextual cueing effect if its time is prolonged. To our knowledge, only two studies have examined the role of initial perception in the contextual cueing effect (Chaumon, Drouet, & Tallon-Baudry, 2008; Zhao et al., 2012). However, Chaumon et al. support the role of perceptual processing and Zhao et al. did not find a role, maybe due to the very short time for perceptual processing. In fact, although Geyer et al. (2010) aimed to examine whether the contextual cueing effect of a pop-out visual search could benefit from other factors except for response selection, to some extent this study suggested that perceptual processing could play a role. Each trial in their study started with the presentation of “placeholder” squares marking the locations of the subsequent search items. In this way, for one thing,

participants had additional “preview” time to process the displays, and for another, they could not actually search for targets. Furthermore, they asked participants to judge if targets exist. The display exposure was time limited or with stress on response speed in their Experiment 2. So, this could rule out later sources contributing to the contextual cueing effect. They found that a contextual cueing effect could exist in pop-out visual search, even without benefits from response selection. Moreover, King, Korb, and Egnér (2012) revealed that learned associations could link closely with the bottom-up information, and together facilitate the visual search task. We consider that the bottom-up information does facilitate the phase of perceptual processing, and the learned associations contribute to attentional guidance. So, we can conclude that perceptual processing is possibly a contributing factor to the contextual cueing effect. However, some studies have suggested that bottom-up information (e.g., colored backgrounds, segmentation by salient features) played little or negative roles in the contextual cueing effect (Conci & Mühlénen, 2009; Kunar, John, & Sweetman, 2014a). Given the above, future studies should further examine the interaction between perception and contextual cueing effect under different conditions.

In low contrast, except for the role of perceptual processing, there was a role of attentional guidance – in fact, it played a major role. In Experiment 1, the benefit of time on the perceptual processing phase in repeated displays was about 8 ms, while the mean number of fixations required to find targets reduced about 0.67 in repeated displays. Similarly, in Experiment 2, the time of perceptual processing saved about 11 ms in repeated displays, and the mean number of fixations saved about 0.52. We defined this as a fixation only, when the fixation duration was no less than 50 ms. So, we can estimate that the time saved due to the decreasing number of fixations in repeated displays was greater compared with the time saved from perceptual processing. Thus, the improved attentional guidance contributes most to the contextual cueing effect, even in low-contrast conditions. These results were consistent with a large body of previous studies (Chun & Jiang, 1998; Harris & Remington, 2017; Peterson & Kramer, 2001; Zhao et al., 2012).

Unexpectedly for us, with low contrast of both experiments, the decision times were longer for repeated than for novel displays, but this cost on response-related processing was not observed with high contrast. We supposed that this resulted from our proxy for response-related processes – decision time. We defined this as the duration between the last eye fixation and key-press response (Zhao et al., 2012). This index was suitable for high-contrast trials, but not for low-contrast trials. With low contrast, the high similarity between item color and

background color made participants compare the target with surrounding distractors repeatedly, and only then could they decide whether it was the target. Consequently, participants must have seen the target before the last fixation. In other words, it was not possible that participants did not see the target until the last fixation. Consistent with this suggestion, we examined the data of participants’ scan path, and found that it was not the last fixation for which participants put their fixation on the target the first time in most trials, in other words, participants fixated the target, moved away, and then returned before making a response on most low-contrast displays. Given the above, we can speculate that our proxy for response-related processes (decision time) underestimated the actual amount of time for response-related processing with low contrast.

In general, whether there is a role of response-related processing in the contextual cueing effect of low contrast remains to be further examined. But at least in high-contrast displays, our study did not observe facilitation from response-related processing. Congruently, Harris and Remington (2017) suggested that even in a parallel-search condition, the contextual cueing effect did not benefit from the phase of response selection. By contrast, Zhao et al. (2012) found that facilitation of response selection also played a role, although the role was small. One possible reason for our high-contrast trials lacking facilitation from response selection was the weak contextual cueing effect (the RTs benefit was only 26 ms). RTs in the study by Zhao et al. (2012) benefitted by around more than 100 ms. Since the total benefit was very small, it was hard to find a statistically significant difference even there was a role of response-related processing. Another possible reason was the difference in stimuli between Zhao et al. (2012) and our study. They used corrected displays, in which the size of each item was equal according to the eccentricity from the fixation point. However, the displays used by us were the most common one, as in Chun and Jiang (1998). As a result, future research needs find out the condition in which response-related processing plays a role.

Similarly, with regard to the controversy between Zhao et al. (2012) and Harris and Remington (2017) about the role of response selection, one of the reasons was also their different stimuli. The stimuli used by Harris and Remington (2017) were the same as ours, the most common ones, as in Chun and Jiang (1998). Maybe corrected displays used by Zhao et al. (2012) made the role of response selection more likely to occur, because the size of each item was equal according to the eccentricity from the fixation point, and participants tend to make less eye movements, leaving more room for response-related processes to play a role.

Another possible reason was the specific parameter, proxy for response-related process. Zhao et al. (2012) defined it as the time between the last eye fixation and response, while Harris and Remington (2017) defined it as the time between when the eyes first fixate the target and the response. It is possible that the last eye fixation is not the first time of it landing on a target, as we found in our data of eye movements, sometimes participants fixated the target, moved away, and then returned before responding. However, the ambiguity in defining the proxy for the response-related process is a problem inherent to all eye-tracking studies, as noted by Sisk, Remington, and Jiang (2019). So, maybe we need to find new methods to examine the role of response-related processing.

It is worth noting that one of our experimental hypotheses may be puzzling or controversial: the amount of contextual cueing effect would be greater in low-contrast than in high-contrast if there is a role of perceptual processing. One could think that the attentional guidance and response-related processing perhaps play different roles in different display contrast, in other words, the greater contextual cueing effect of low-contrast displays may be due to other factors (such as attentional guidance, response-related process) instead of perceptual processing. Indeed, we think that it is possible according to our eye-movement data, but we could not know this before our experiments. In addition, before implementation of the experiments, we had preconceived that even if this is the case, we still can use our eye-movement data (initial saccade latency) to examine whether initial perceptual processing plays a role in the contextual cueing effect.

In sum, this study validates the role of perceptual processing in the contextual cueing effect as long as the perceptual processing time is prolonged, moreover, in favor of the attentional guidance hypothesis. As a whole, we speculate that whether the contextual cueing effect benefits from improved perceptual processing and/or attentional guidance and/or improved response-related processing, depends on the time during which the corresponding phase proceeds. If a specific phase needs more time to proceed, the corresponding factor will play a larger role. This needs further investigation in future studies.

Acknowledgments This research was partly supported by Shandong Provincial Key Research & Development Program (No. 2018GSF118090), Natural Science Foundation of Shandong Province (No. ZR2017MC058) and a Visiting Scholar Grant from Shandong Provincial Government (No. 2017190) to Y. R. We are grateful to Professor Richard A. Abrams (Department of Psychological and Brain Sciences, Washington University in St. Louis) for helpful discussions of earlier drafts of this paper. We also thank the Editor Professor/Dr. Michael

D. Dodd (Department of Psychology, University of Nebraska-Lincoln), Dr. Anthony Harris, and another two anonymous reviewers for their helpful comments and suggestions on earlier versions of the manuscript.

Open Practices Statement The behavioral and eyetracking raw data will be made available upon request.

Author Contributions F. Zhao and Y. Ren conceived and designed the experiments. F. Zhao performed the experiments under the supervision of Y. Ren. F. Zhao analyzed the data. F. Zhao and Y. Ren wrote the manuscript.

Compliance with ethical standards

Conflicts of interest All authors declare that they have no conflicts of interest. The authors declare no competing financial interests.

References

- Chaumon, M., Drouet, V., & Tallon-Baudry, C. (2008). Unconscious associative memory affects visual processing before 100 ms. *Journal of Vision*, 8(3), 10. <https://doi.org/10.1167/8.3.10>
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36(1), 28–71.
- Conci, M., & Mühlhelen, A. (2009). Region segmentation and contextual cuing in visual search. *Attention, Perception, & Psychophysics*, 71(7), 1514–1524.
- Geringswald, F., Baumgartner, F., & Pollmann, S. (2012). Simulated loss of foveal vision eliminates visual search advantage in repeated displays. *Frontiers in Human Neuroscience*, 6, 134. <https://doi.org/10.3389/fnhum.2012.00134>
- Geyer, T., Zehetleitner, M., & Müller, H. J. (2010). Contextual cueing of pop-out visual search: When context guides the deployment of attention. *Journal of Vision*, 10. <https://doi.org/10.1167/10.5.20>
- Harris, A. M., & Remington, R. W. (2017). Contextual cueing improves attentional guidance, even when guidance is supposedly optimal. *Journal of Experimental Psychology: Human Perception and Performance*, 43(5), 926–940.
- Higuchi, Y., & Saiki, J. (2017). Implicit learning of spatial configuration occurs without eye movement. *Japanese Psychological Research*, 59(2), 122–132.
- Hout, M. C., & Goldinger, S. D. (2012). Incidental learning speeds visual search by lowering response thresholds, not by improving efficiency: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 38(1), 90–112.
- Johnson, J. S., Woodman, G. F., Braun, E., & Luck, S. J. (2007). Implicit memory influences the allocation of attention in visual cortex. *Psychonomic Bulletin & Review*, 14(5), 834–837.
- Kasper, R. W., Grafton, S. T., Eckstein, M. P., & Giesbrecht, B. (2015). Multimodal neuroimaging evidence linking memory and attention systems during visual search cued by context. *Annals of the New York Academy of Sciences*, 1393(1), 72–73.
- King, J. A., Korb, F. M., & Egner, T. (2012). Priming of control: Implicit contextual cuing of top-down attentional set. *Journal of Neuroscience*, 32(24), 8192–8200.
- Kojima, H., & Kawabata, Y. (2012). Perceived duration of chromatic and achromatic light. *Vision Research*, 53(1), 21–29.
- Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2006). Contextual cueing by global features. *Perception & Psychophysics*, 68(7), 1204–1216.
- Kunar, M. A., Flusberg, S. J., Horowitz, T. S., & Wolfe, J. M. (2007). Does contextual cueing guide the deployment of attention? *Journal*

- of *Experimental Psychology: Human Perception and Performance*, 33(4), 816–828.
- Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2008). Time to guide: Evidence for delayed attentional guidance in contextual cueing. *Visual Cognition*, 16(6), 804–825.
- Kunar, M. A., John, R., & Sweetman, H. (2014). A Configural Dominant Account of Contextual Cueing: Configural Cues are Stronger than Colour Cues. *Quarterly Journal of Experimental Psychology*, 67(7), 1366–1382.
- Kunar, M. A., Watson, D. G., Cole, L., & Cox, A. (2014). Negative Emotional Stimuli Reduce Contextual Cueing but not Response Times in Inefficient Search. *Quarterly Journal of Experimental Psychology*, 67(2), 377–393.
- Manelis, A., & Reeder, L. M. (2012). Procedural learning and associative memory mechanisms contribute to contextual cueing: Evidence from fMRI and eye-tracking. *Learning & Memory*, 19(11), 527–534.
- Meyer, T., Quaedflieg, C., Bisby, J.A., & Smeets, T. (2019). Acute stress – but not aversive scene content – impairs spatial configuration learning. *Cognition and Emotion*. <https://doi.org/10.1080/02699931.2019.1604320>
- Musz, E., Weber, M.J., & Thompson-Schill, S. L. (2015). Visual statistical learning is not reliably modulated by selective attention to isolated events. *Attention, Perception, & Psychophysics*, 77(1), 78–96.
- Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Attention, Perception & Psychophysics*, 63(7), 1239–1249.
- Pollmann, S. (2019). Working memory dependence of spatial contextual cueing for visual search. *British Journal of Psychology*, 110(2), 372–380. <https://doi.org/10.1111/bjop.12311>
- Schankin, A., Hagemann, D., & Schubö, A. (2011). Is contextual cueing more than the guidance of visual-spatial attention? *Biological Psychology*, 87(1), 58–65.
- Schankin, A., & Schubö, A. (2009). Cognitive processes facilitated by contextual cueing: Evidence from event-related brain potentials. *Psychophysiology*, 46(3), 668–679.
- Schankin, A., & Schubö, A. (2010). Contextual cueing effects despite spatially cued target locations. *Psychophysiology*, 47(4), 717–727.
- Schlagbauer, B., Rausch, M., Zehetleitner, M., Müller, H. J., & Geyer, T. (2018). Contextual cueing of visual search is associated with greater subjective experience of the search display configuration. *Neuroscience of Consciousness*, 2018(1), niy001. <https://doi.org/10.1093/nc/niy001>
- Sewell, D. K., Colagiuri, B., & Livesey, E. J. (2017). Response time modeling reveals multiple contextual cuing mechanisms. *Psychonomic Bulletin & Review*, 25(6), 1644–1665.
- Sisk, C. A., Remington, R. W., & Jiang, Y. V. (2019). Mechanisms of contextual cueing: A tutorial review. *Attention, Perception, & Psychophysics*, 81(8), 2571–2589.
- Thiessen, E.D., Kronstein, A.T., & Hufnagle, D.G. (2013). The extraction and integration framework: a two-process account of statistical learning. *Psychological Bulletin*, 139(4), 792–814.
- Vadillo, M. A., Konstantinidis, E., Shanks, D. R. (2016) Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review* 23 (1), 87–102.
- Zang, X., Jia, L., Müller, H. J., & Shi, Z. (2015). Invariant spatial context is learned but not retrieved in gaze-contingent tunnel-view search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 807–819.
- Zhao, G., Liu, Q., Jiao, J., Zhou, P., Li, H., & Sun, H.-j. (2012). Dual-state modulation of the contextual cueing effect: Evidence from eye movement recordings. *Journal of Vision*, 12(6), 11. <https://doi.org/10.1167/12.6.11>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.