Dual-state modulation of the contextual cueing effect: Evidence from eye movement recordings

Guang Zhao
Key Laboratory of Cognition and Personality, Ministry of Education, Southwest University, Chongqing, China; School of Psychology, Southwest University, Chongqing, China

Qiang Liu
Research Center of Psychological Development and Education, Liaoning Normal University, Dalian, China

Jun Jiao
Key Laboratory of Cognition and Personality, Ministry of Education, Southwest University, Chongqing, China; School of Psychology, Southwest University, Chongqing, China

Peiling Zhou
Key Laboratory of Cognition and Personality, Ministry of Education, Southwest University, Chongqing, China; School of Psychology, Southwest University, Chongqing, China

Hong Li
Research Center of Psychological Development and Education, Liaoning Normal University, Dalian, China

Hong-jin Sun
Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Ontario, Canada

The repeated configurations of random elements induce a better search performance than that of the displays of novel random configurations. The mechanism of such contextual cueing effect has been investigated through the use of the RT × Set Size function. There are divergent views on whether the contextual cueing effect is driven by attentional guidance or facilitation of initial perceptual processing or response selection. To explore this question, we used eye movement recording in this study, which offers information about the substages of the search task. The results suggest that the contextual cueing effect is contributed mainly by attentional guidance, and facilitation of response selection also plays a role.

Keywords: visual search, attentional guidance, response selection, eye movement


Introduction

Previous research has shown that participants respond faster to a target if it is always presented at the same location and accompanied by the same contextual configuration than if it is not. The effect that invariant configuration of the same arrangement of random elements (repeated or predictive display) induces high-grade searching-task efficiency has come to be known as the contextual cueing effect (Brady & Chun, 2007; Chua & Chun, 2003; Chun & Jiang, 1998; Chun & Phelps, 1999; Kawahara, 2003; Kawahara, 2007; Kunar, Flusberg, & Wolfe, 2006; Ogawa & Kumada, 2008; Olson, Chun, & Allison, 2001). It has been suggested that contextual cueing is mediated by implicit memory processes that allow observers to acquire useful information about the structure of the visual context without explicit awareness of the knowledge learned (Brady & Chun, 2007; Chaumon, Drouet, & Tallon-Baudry, 2008; Chua & Chun, 2003; Chun & Jiang, 1998; Chun & Jiang, 1999; Chun &
In the classic contextual cueing experiment, observers were required to respond to the orientation of a T shape (presented in left and right orientations) among a series of L shapes, which were rotated randomly (Chun & Jiang, 1998). Although the layouts of the displays (and target locations) were different from trial to trial, half of the displays repeated across blocks throughout the experiment session. Participants were able to respond to the target faster for those repeated displays. To reveal the mechanism of the contextual cueing effect, Chun and Jiang (1998) varied the number of distracter items (set size), then fitted a line to the RT × Set Size function, comparing the search slopes and intercepts in repeated contexts with those in the nonrepeated contexts. The slope of the fitted line was taken as a measure of search efficiency to reflect the cost of adding an item to the search display, whereas nonsearch factors such as initial perceptual processing and response selection processing were believed to contribute to the intercept. If the contextual cueing effect was the result of learning to direct attention to the target, there would be a downward trend in the search slope over the course of the experiment, and such a downward trend should be more prominent for repeated display. In contrast, if contextual cueing was accounted for by perceptual-recognition processing or response selection, the intercepts would decrease over time and more so for the repeated displays (Figure 1a).

Empirical results in the literature did not reach an agreement on which stage of the process is facilitated by the contextual cueing in predictive displays. Some studies have demonstrated that, with repeated configurations, the fitting function showed a more efficient searching performance, supported by a lower and steeper search slope rather than the intercept (Alvarez, Konkle, & Oliva, 2007; Chun & Jiang, 1998; Crewther, Lawson, & Crewther, 2007; Duncan & Humphreys, 1989; Rausei, Makovski, & Jiang, 2007; Wolfe, 1998). These results confirmed that contextual cueing was driven by the deployment of attentional guidance.

However, in the study of Kunar et al. (Kunar, Flusberg, Horowitz, & Wolfe, 2007), significant decrease in slope was not found in the predictive/random comparison, but differences in the intercepts were detected. Furthermore, once interference was manipulated in the response selection process, the contextual cueing effect would disappear. Accordingly, Kunar et al. proposed that repeated configurations may reduce the threshold needed to respond to the target; therefore, contextual cueing acts, at least in part, by speeding responses to targets in a familiar context (Kunar et al., 2007).

The discrepancy in results between different studies suggests that the slope and intercept of the fitting line of RT might not be the most effective way in revealing the mechanisms of the contextual cueing effect. One possible reason for the discrepancy in results between studies is that the search slope and intercept were obtained based on the overall response time. However, overall response times might not be the most sensitive measure to determine the factors that affect the steepness of the slopes (Peterson & Kramer, 2001; Tseng & Li, 2004). In addition, it is generally believed that the decrease in the search slope should be attributed to the reduction of the averaged time per search item. However, studies have confirmed that, during searching processes in predictive configurations, participants might only need to search through a few visual items before accurately locating the target’s position (Peterson & Kramer, 2001, Tseng & Li, 2004).

It is well known that the eye tracking system can be used to infer moment-to-moment cognitive processes in a fairly direct manner. Hence, by means of eye movement recording, contextual cueing can be better investigated separately for different processing stages (Chun & Jiang, 1998). Though several studies have introduced eye movement to investigate mechanisms of the contextual cueing effect, they varied in stimulus used and the data analysis. Thus, these studies cannot fully resolve which process stage(s) would contribute to the contextual cueing effect. For example, Peterson and Kramer (2001) were the first to investigate whether recognition and attentional guidance can modulate contextual cueing using the eye tracking method. Their results showed a reduction of the number of fixations that accompanies the reduction of search times in repeated displays. They did not investigate the process of response selection. Tseng and Li (2004) also used the eye tracking method and found the effect of attentional guidance. No facilitation of response selection in contextual cueing displays was found as indicated by the lack of differences in the time between last fixation to button press (TLF to BP). It is important to note that Tseng and Li (2004) used a different type of display. Instead of using the classic contextual cueing paradigm in which the contextual cue was composed by a layout of the target T and distractor Ls, Tseng and Li (2004) introduced a display by which the contextual cue was available from another feature (a layout of blue disks) in the scene that physically did not resemble T or L. Because their eye movement results revealed that subjects did not exert the eye fixations on cue elements, the mechanism revealed in this paradigm might not apply to the more commonly used paradigm in which the target/distractors themselves form the familiar context.

We thus adopted a task paradigm that resembles the classical and more commonly used paradigm to explore how contextual cueing was elicited. In addition to RT, we explored oculomotor correlates of contextual cueing...
effects based on the behavioral discrepancy revealed in RT × Set Size function. Further, considering that the contextual facilitation of response selection was discovered by Kunar et al. (2007), we used a stimulus similar to that introduced by Kunar in order to have a better opportunity to explore the phase of response selection. We selected two sets of parameters (Figure 1b) that might reflect the nature of eye movements in this task: (1) variables for the duration of different stages of the responses and (2) variables for the eye position and the saccade sequence.

First, we analyzed the time course of the search task by recording two variables: initial saccade latency (IL) and the time between last eye fixation and response through the button press (TLF to BP). *Initial latency* is the period from the onset of the display to the initiation of the saccade (Nakatani & Pollatsek, 2004; Rayner, 1998; Tseng & Li, 2004). Although this period is short, there is converging evidence showing that initial latency can be influenced by a process of perceptual recognition. TLF to BP is the variable that would presumably measure the reaction time required for a participant to respond when the fixation is close to the target location (Tseng & Li, 2004). It was suggested that this duration could be related to a decrease in the RTs of the repeated contexts (Kunar et al., 2007; Tseng & Li, 2004). Consequently, this parameter could be used to reflect the exact features processed in the response selection stage.

After obtaining the two parameters previously described, we partitioned the entire response RTs into three segments—the early phase, the middle phase, and the late phase. The early phase corresponds to the initial latency, which could be related to initial perceptual processing. The late phase corresponds to TLF to BP, which could be related to response selection. The duration of the middle phase is the remaining duration (obtained by subtracting the initial latency and TLF to BP from the overall manual response RT) which could be related to attentional guidance. We analyzed the duration for these three phases separately, which allowed us to identify the magnitude of the contextual cueing effect in each phase.
The second set of the eye movement variables include distance between first eye fixation and the target, D(first fix, T) and scan pattern ratio (SPR), which is obtained by dividing the total distances traveled by all the eye movements prior to arriving at the target by the linear distance between the first fixation point and the target. This variable reflects how directly the eyes move to the target (Brockmole & Henderson, 2006; Henderson, Weeks, & Hollingworth, 1999). Repeated contexts can also modulate the guidance of attention to known target positions (Brockmole & Henderson, 2006; Peterson & Kramer, 2001; Tseng & Li, 2004). Therefore, in the visual search task, both these distance parameters can be used to reflect how observers process visual input in the attentional guidance stage.

We also analyzed average fixation duration and number of saccades. Fast search performance could conceivably be caused by shorter average fixation duration or smaller number of saccades/fixations or both. Studies have showed that, in repeated configuration contexts, participants needed fewer fixations to locate the target (Peterson & Kramer, 2001; Torralba, Oliva, Castelhano, & Henderson, 2006). Tseng and Li (2004) identified the number of saccades as the index for the contextual effect.

In summary, in this study, through comparison of manual response and eye movement variables, we hope to offer an in-depth and comprehensive analysis of the underlying process of the contextual cueing effect. Through the analysis of the duration of the different stages of the process and various eye movement parameters, we also performed a more direct exploration of the modulation in the initial perceptual process, attentional guidance, and response selection by analyzing the duration of each stage as well as various eye movement parameters. Note that our second set of variables describes events in the middle phase, as they characterize different aspects of the same eye movement. In the present study, we showed that the middle phase could be a major contributor to the contextual cueing effect; therefore, it is useful to reveal different characteristics of the same oculomotor events in this phase of the task.

Materials and methods

Participants

Thirty-eight undergraduates (15 males and 23 females) were paid to participate in the experiment. All had normal or corrected-to-normal vision. None of the participants were aware of the purpose of the study nor had they participated in a visual search experiment before.

Apparatus and stimuli

The experiment was carried out in a quiet and isolated room with a dark curtain to separate participants and the experimenter from each other. The participants were instructed to sit at a 60-cm distance from a 19-inch CRT display monitor with a sample rate of 85 Hz and to press a keyboard in response to stimuli. Eye movements were recorded with an Eyelink tracker (EyelinkII, SR Research, Toronto) with 250-Hz temporal resolution and a 0.2° spatial resolution. An infrared tracking system tracked head motion.

Stimuli were generated using a Matlab program. All the items in the stimulus consisted of two lines of equal length (forming either an L or a T shape). The background color of the screen was uniform gray. Three black concentric circles surrounded the fixation point with diameters of 9.5°, 15.5°, and 25° visual angle. Sixteen black lines radiated from the fixation point and were roughly equidistant from one another forming a radial lattice. On every trial, either 8 or 12 (depending on the set size) circular placeholders appeared at the conjunctions between the concentric circles and the spokes. To explore whether contextual information can provide facilitation before the first saccade, we corrected the size of each letter based on the eccentricity from the fixation point to ensure that each item had an equal visibility during that phase. Specifically, the size of the place-holding circles and the Ts and Ls were made to be proportional to the eccentricity. Accordingly, the diameters of the placeholders were varied from the fixation point to project outward at 2°, 3.3°, and 5.4°, and the diameters of the stimulus items correspondingly subtended visual angles of 1° × 1°, 1.5° × 1.5°, and 2.5° × 2.5°, respectively. To rule out location probability effects, the target appeared equally often in each of 16 possible locations throughout the experiment—eight for predictive configurations and eight for random configurations. Meanwhile, the 16 positions were roughly equally distributed among the three concentric circles and four quadrants of the stimulus.

The target T stimulus was rotated by 90° to the left or right, while the distractor stimuli of the Ls were presented randomly in one of four orientations (0°, 90°, 180°, and 270°). The participants were asked to respond to the direction of the target T among the distractors. An example stimulus is illustrated in Figure 2.

The stimulus was similar to Kunar’s (Kunar et al., 2007), except that we did not manipulate the color attribute of the placeholders and stimuli. Our reasoning was that, because the contextual cueing effect was
driven by the association between the global scene configuration and the position of the targets, the effect can be enhanced by limiting the local cues.

Design and procedure

All the configurations were divided into two conditions—predictive configurations and random configurations. Each random configuration was only presented once in the experiment, while the predictive sets of trials were repeated across blocks throughout the entire experiment.

The entire session included 28 blocks of 16 trials, and each block contained eight random and eight predictive trials. Across the blocks, the predictive trials had fixed placeholder configurations. Two different set sizes (8 and 12) were tested in the experiment and were randomized within the blocks. In order to increase the power of our statistical analyses, all four adjoining blocks were grouped into one epoch, resulting in a total of seven epochs.

The entire experiment consisted of two main sessions: one was a visual search task, and the other was a recognition test. In the visual search task, each trial started with a fixation display. The participants were instructed to sit facing the monitor, keeping their body and head steady while staring into the fixation point. The experimenter pressed the spacebar on a spare keyboard to begin the trial. A fixation display lasted for 500 ms and was then replaced by the search display. Participants searched for the target and pressed a key immediately upon detection. They pressed the left “control” key if the target was pointing left and the right “control” key if it was pointing right. The trial terminated if no response was given within 10 s. Immediately after the participant responded, the display was cleared and replaced by a blank gray screen. After a brief duration of 200 ms, a “next” screen was displayed for the next trial.

After the visual search task, the participants performed a recognition test that was not revealed in advance. They were asked to respond if they noticed some configurations that were repeated from block to block. This task had no time limit. For a positive response, they pressed the number “1,” and for a negative response, they pressed number “2” on the keyboard. All the configurations in this procedure included 24 displays, divided into three parts: eight configurations formed a predictive set, eight formed a random set, and eight configurations had not been exposed previously. All the configurations were balanced in a set size.

Eye movement measures

Initial latency was the time elapsed between display onset and initiation of the first saccade (Nakatani & Pollatsek, 2004; Rayner, 1998). Saccades and fixations were defined using the saccade detection algorithm supplied by SR Research; saccades were identified by deflections in eye position in excess of 0.1° with a minimum velocity of 30°/s and a minimum acceleration of 8000°/s². We partitioned saccades into ineffective and effective saccades by adopting Tseng and Li’s (2004) calculation of the onset of the effective phase. We regarded ineffective and effective saccades as ineffective and effective search phases, respectively. For each trial, an effective saccade was a saccade that brought the eye to the largest distance to the target. We assume any saccades including and after that saccade were a valid search saccade (start of a effective search phase).

D(first fix, T) is the distance between the first eye fixation and the target after display onset. SPR is calculated by taking the length of the scan pattern taken by the participants’ eyes through the display on his or her way to the target object (computed as the summed distance between all fixations from display onset to the first fixation on the target) divided by the most direct possible path (computed as the distance from the central fixation point to the center of the target object) (Brockmole & Henderson, 2006; Henderson, Weeks, & Hollingworth, 1999).

TLF to BP is the time from last fixation to button press. The “last eye fixation” was defined as one of the last two fixations that was spatially closer to the target. This qualification was required because it was assumed that target identification occurred at the point when eye fixation was closest to the target, while for a very small number of trials, the chronologically last fixation...
brought the eye away from the target. In addition, a minimum of 150 ms was deemed to be required of TLF to BP. An eye fixation that resulted in a TLF to BP of less than 150 ms was disqualified and replaced by its predecessor (Tseng & Li, 2004).

Results

Behavioral responses

Mean accuracy in the explicit recognition task was 48%. Participants correctly classified predictive displays as predictive on 50% of the trials (hit rate) and classified random displays and not-exposed displays as non-predictive on 48% and 44%, respectively (false alarm rate). However, they did not differ significantly among them ($F[2, 110] = 0.610$, $p = 0.552$). These findings suggest that participants were not aware of the predictive configurations, which is consistent with what was found in Chun and Jiang’s study (1998). For the visual search task, as the overall error rates were less than 0.6% in both predictive and random conditions, we thus focused only on the RT data in the data analysis. We excluded trials (less than 0.51% of all trials) for which the RTs were below 200 ms or above 4000 ms.

The mean RT values for each configuration condition as a function of epoch for set sizes of 8 and 12 are plotted in Figure 3a (left and right panels, respectively). For RT data, a $2\times7\times2$ repeated measures ANOVA (Table 1a) demonstrated significant main effects of configuration, epoch, and set size. Significant two-way interactions were obtained for configuration $\times$ epoch, indicating a greater downtrend for predictive configurations as the epoch session progressed; set size $\times$ configuration, showing a greater benefit for predictive displays for the larger set size; and set size $\times$ epoch, indicating a more efficient search for the larger set size as the session progressed. Moreover, the three-way interaction between configuration $\times$ epoch $\times$ set size was also significant, indicating a more efficient search for the predictive displays and for the larger set size as the session progressed.

The search slopes and intercepts are plotted as a function of the epoch in Figure 4 (left and right panels, respectively). For the repeated measures ANOVA (Table 1b) on the slope data, we observed a main effect of configuration and epoch, and there was also a significant interaction between configuration and epoch. These findings suggest that attentional guidance likely plays an important part in the contextual cueing effect. To quantify the size of the effect of configuration, we collapsed the data across the last three epochs (epochs 5 to 7); the average slope for the predictive display was 31.71 ms/item. The gain time for predictive trials was 28.31 ms/item faster than those for random trials ($t[37] = 2.828$, $p = 0.007$). For the intercepts, the results revealed no significant main effects of configuration, epoch, or interaction. Furthermore, intercepts for the predictive displays fluctuated at about 800 ms across the epochs, demonstrating no decreased tendency.

Data from the eye movement recording, part 1: duration of the three phases

We first partitioned the entire duration of the response into three phases (early, middle, and late phases). Figure 3b, 3c, and 3d plot the duration of the three phases as a function of epoch for set size of 8 (left panel) and 12 (right panel). To analyze each phase, we entered the configuration, epoch, and set size into three factors repeated measures ANOVA with results showed in Table 1a. For the duration of the early phase, no significant main effects or interactions were found. For the duration of the middle phase, we observed significant main effects of configuration, epoch, and set size.
Significant two-way interactions were obtained for configuration × epoch, indicating a greater downward trend for predictive configuration as the epoch session progressed; for set size × configuration, showing a greater benefit for the larger set size; and for set size × epoch, indicating a more efficient search for the larger set size as the epoch session progressed. The three-way interaction between configuration × epoch × set size was marginally significant.

For the duration of the late phase, we observed significant main effects of configuration, set size, and epoch. Neither two-way nor three-way interactions were significant. Although the configuration × epoch interaction was not significant, we also analyzed the data in the first epoch by comparing the data in the first four blocks. The results showed a general trend of interaction for configuration × epoch × set size was marginally significant.

For the duration of the late phase, we observed significant main effects of configuration, set size, and epoch. Neither two-way nor three-way interactions were significant. Although the configuration × epoch interaction was not significant, we also analyzed the data in the first epoch by comparing the data in the first four blocks. The results showed a general trend of interaction for configuration × epoch × set size was marginally significant.

Therefore, the data from the group of 38 participants reveals that the contextual cueing effect can be manifested mainly in the middle phase and to some extent in the late phase.

Data from the eye movement recording, part 2: ocular parameters

Considering that the middle phase (search phase) could be one of the major contributors to the overall contextual cueing effect revealed through RT, we next examined four eye movement parameters—D(first fix,T), scan pattern ratio, averaged fixation duration, and the saccade number—to identify which aspects of the eye movement would explain the contextual cueing effect. In Figure 6, we first plotted the averaged value for D(first fix,T), scan pattern ratio, and averaged fixation duration as a function of epoch for set size of 8 (left panel) and 12 (right panel), respectively. A 2(configuration) × 7(epoch) × 2(set size) repeated measures ANOVA were conducted for each of the three variables, and the results are shown in Table 2.

For D(first fix,T), we only observed significant main effects of epoch, and no other main effects or interactions were significant. For the scan pattern ratio, we observed significant main effects of configuration, epoch, and set size. Significant two-way interactions were also obtained for epoch × configuration and epoch × set size. Interaction for configuration and set size was not significant, indicating that the slope for the predictive and random display would not be

![Image](image-url)

**Figure 4.** Slopes (left panel) and intercepts (right panel) of total response RT as a function of epoch for random and predictive configurations.

**Figure 5.** The mean duration for the late phase as a function of blocks (in the first epoch) in each configuration for set size 8 (left panel) and set size 12 (right panel).
significantly different, suggesting little contribution of scan pattern ratio to the attentional guidance revealed in the duration of the middle phase and overall RT. For the averaged fixation duration, none of the main effects or interactions were significant. The average fixation duration tended to remain in a narrow range of 220 to 230 ms.

If the averaged duration for each fixation does not fluctuate much, given that there was a great variation of middle (search) phase durations, the total number of fixations (or saccades) might be the major contributor to the overall difference in duration. We thus analyzed the data for the number of saccades. The analysis of the total number of saccades (Figure 7a and Table 3) revealed significant main effects of configuration, epoch, and set size. Significant two-way interactions were also obtained for configuration · epoch, showing a greater downtrend of saccade numbers for predictive configuration as the epoch session progressed, and set size · epoch, indicating more efficient search for the larger set size as the epoch session progressed. The two-way interaction between set size × configuration was not significant, indicating that, after collapsing data across seven epochs, the slope for the predictive and random display would not be significantly different.

The nonsignificant slope result for the total number of saccades does not necessarily mean that the number of saccades is not related to the contextual cueing effect. During visual search, participants could conceivably first make some saccades that involve a random scan of the scene, then after some form of recognition of the scene, they start to “zoom” into a certain region of the scene. Based on the method described by Tseng and Li (2004), we partitioned the series of saccades within one trial into two subphases. In the first phase, the ineffective search phase (consisting of ineffective saccades, Figure 7b), the eye movement is more or less a random movement. Then, in the second phase, the effective search phase (consisting of effective saccades, Figure 7c), the eye movement brings the fixations increasingly closer to the target.

For the ineffective search phase, ANOVA results (Table 3) reveal that all the main effects and interactions were significant, including the interaction between configuration and set size, indicating that the slope for the predictive and random display was also significantly different. However, for effective search phase, all the main effects and interactions were significant except for the interaction between configuration and set size.

Figure 8 shows that the calculated slopes for the total number of saccades: ineffective and effective saccades. To quantify the size of the effect of different types of saccades, we collapsed the data across the last three epochs (epochs 5 to 7). The gain (the number of saccades for random display minus that for predictive displays) was 0.026 for total saccades, 0.019 for ineffective saccades, and 0.0074 for the effective saccades. This indicates that reduction in number of ineffective saccades for the predictive display contributed to 72% of the reduction in total number of saccades, while the effective saccades contributed to 28%.

<table>
<thead>
<tr>
<th></th>
<th>D(first fix, T)</th>
<th>Scan pattern ratio</th>
<th>Fixation duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configuration</td>
<td>0.367</td>
<td>38.464***</td>
<td>2.225</td>
</tr>
<tr>
<td>Epoch (F[6, 222])</td>
<td>8.595***</td>
<td>44.205***</td>
<td>1.070</td>
</tr>
<tr>
<td>Set size (F[1, 37])</td>
<td>3.180</td>
<td>203.302***</td>
<td>2.178</td>
</tr>
<tr>
<td>Configuration × Epoch (F[6, 222])</td>
<td>1.717</td>
<td>2.517**</td>
<td>0.433</td>
</tr>
<tr>
<td>Configuration × Set size (F[1, 37])</td>
<td>0.018</td>
<td>2.182</td>
<td>1.639</td>
</tr>
<tr>
<td>Epoch × Set size (F[6, 222])</td>
<td>1.510</td>
<td>4.230***</td>
<td>1.558</td>
</tr>
<tr>
<td>Configuration × Epoch × Set size (F[6, 222])</td>
<td>1.463</td>
<td>1.218</td>
<td>0.152</td>
</tr>
</tbody>
</table>

Table 2. ANOVAs of three eye movement parameters.
In addition to the general pattern of data in a population, we also explored the data by examining the distribution of different types of saccades in individual trials. Figure 9 plots the distribution of the trials that involved different types of saccades for predictive and random configurations for set size 8 (left panel) and 12 (right panel), respectively, and for the first two epochs (Figure 9a) and last two epochs (Figure 9b), respectively. In general, for the total number of saccades and effective saccades, the distribution peak was at 3 to 5, indicating that for a great proportion of the trials, three to five eye movements were required before the eye landed at the target region. For ineffective saccades, the distribution peak was at zero (more than 40% of the trials), indicating that for more than 40% of the trials, the eye moved closer to the target starting at the first saccade. For about 20% of the trials, the eye first moved away from the target before moving closer to the target in the second saccade.

The comparison of Figure 9a and 9b reveals the dynamics of the proportion change of different types of saccades at the beginning and later phases of learning. For the first two epochs, the distributions for predictive and random displays mostly overlap; however, in the last two epochs, the distribution shifts leftward, indicating smaller numbers of saccades were needed as the participants learned the task, and the predictive displays led to greater amount of shift. Consequently, predictive displays led to a greater proportion of a smaller number of saccades and a smaller proportion of a larger number of saccades.

**Discussion**

In the present study, we used eye movement to explore whether the contextual cueing effect was related to attentional guidance, initial perceptual process, late response selection, or any combinations of these mechanisms. We have demonstrated a significant contextual cueing effect in the manual RT measure and a significant difference between predictive and
random displays in search slope, suggesting the role of attentional guidance in the contextual cueing effect. We further investigated the contextual cueing effect through partitioning the RTs into three phases—the early phase, the middle phase, and the late phase. We next specifically examined some ocular indices combined with different types of saccades to explore how context facilitates search.

The results from different processing stages demonstrated that less time was required on the predictive configurations in both the middle phase and the late phase. During the middle phase, the reduced total number of saccades (mostly a reduced number of ineffective saccades) contributed to the overall contextual cueing effect. These results suggest that both attention guidance and response priming could contribute to the contextual cueing effect. These results are next discussed in detail.

First of all, we showed an overall contextual cueing effect in the RT measure. Through the benefit of search slope in the predictive displays, we demonstrated that attentional guidance could be contributing to the overall contextual cueing effect. This is consistent with the results from Chun and Jiang (1998). However, it is notable that our experimental paradigm was very similar to the paradigm used by Kunar et al. (2007), but they did not find or found weak benefit in the search slope. A possible reason for this difference is that the benefit in search slope could be better revealed in a more difficult search task. In Kunar et al.’s (2007) study, different distracter items and targets were of distinctly different colors, while we used a black and white display. The color attributes could facilitate spatial memory, resulting in a faster response. For example, the RT reported in Kunar et al. (2007, e.g., figure 2) tended to be 200 to 300 ms shorter than what we found, thus, the benefit in the search slope was not

Figure 9. The distributions of trials with different numbers of saccades in predictive and random configurations for set size 8 (left panel) and set size 12 (right panel) in the first two epochs (top three panels) and the last two epochs (bottom three panels).
observed in Kunar et al.’s (2007) study. Indeed, in a study by Kunar, Flusberg, and Wolfe (2008) in which a complex background display was used to increase the time taken to find the target, they found a greater guidance benefit in the contextual cueing effect. They suggested that guidance would become available in a contextual cueing task if it was given time to develop (Kunar, Flusberg, & Wolfe, 2008).

In what we define as the early phase of the search (from the onset of stimulus to the onset of saccade), the current results did not reveal the contextual cueing effect in duration. When we consider the data from all the trials, the D(first fix, T) was not significantly different between the predictive display and random display. Therefore, taking the population behavior as a whole, the results reveal little involvement in contextual cueing in the first phase. However, a slightly higher proportion (e.g., about 12% for a set size of 12) of the trials with zero ineffective saccades was observed (Figure 9b), indicating that, for the predictive displays, there were more trials in which the eye moved closer to the target, even in the first saccade. It is possible that, after a short period of latency of about 200 ms, at least for some trials, there were some forms of the guidance of spatial attention toward the target by the repeated configuration. These findings were consistent with the conclusions from previous electrophysiological studies (Johnson et al., 2007; Olson et al., 2001). These studies showed that the context could enhance the initial site of visual cortical processing, including V1, V2, and other portions of the extrastriate cortex, and attentional guidance could increase the amplitude of the N2pc in repeated context, suggesting a greater early allocation of attention to the visual target.

The contextual cueing effect was prominent in the duration of the middle phase of our task, in which most of the eye movement occurs. This result suggests that the visual context in repeated displays can be used to guide attention toward the target locations. In fact, several investigations, using RT measure, have evidenced that attention-directed processing was an important source of contextual cueing (Chun & Jiang, 1998; Chun & Jiang, 1999; Chun & Phelps, 1999; Jiang & Chun, 2001; Kawahara, 2003; Olson et al., 2001; Ono, Kawahara, & Jiang, 2005; Peterson & Kramer, 2001).

The analysis of the eye movement parameters revealed that, among four different ocular parameters, a reduced number of saccades in the predictive display was the main contributing factor in the contextual cueing. On average, the eye movement did not bring the eye significantly closer to the target from the start, the distances of the scan path were not shorter, and participants did not make shorter fixations. Instead, they simply made fewer saccades for the predictive display. The involvement of the number of saccades in the overall contextual cueing effect was also reported by Peterson and Kramer (2001) and Tseng and Li (2004). Moreover, we also found that the reduction of the number of saccades was mostly (72%, calculated based on the gain) caused by a reduction of the number of ineffective saccades (those earlier saccades that did not bring the fixation closer to the target). Reduction in the number of effective saccades contributed to 28% of the reduction in the total number of saccades. These results were generally in line with the finding of Tseng and Li (2004). Using a different way to calculate the contribution of the ineffective and effective saccades, Tseng and Li (2004) found that a reduction of number of ineffective saccades was the only factor in causing the reduction of the total number of saccades.

For the late phase of the response, a reliable effect of configuration on TLF to BP was found, which indicated a shorter duration of TLF to BP or a quicker response to targets in predictive contexts. Considering that the size of the effect of response selection here is relatively small, we collapsed the data across the last three epochs (epochs 5 to 7), and the size of the contextual cueing (the difference between random and predictive display) was about 54 ms and 60 ms for set size 8 and 12, respectively. Overall, these results suggest that the contextual cueing effect, to a small extent, can be contributed by the modulation of the process of response selection. Although behavioral results failed to obtain differences from intercepts between predictive and random configurations, we obtained a reliable contextual facilitation of response selection from eye movement recording, which supported the study by Kunar et al.

In conclusion, the present experiment demonstrates a clear benefit of using eye movement data in identifying the precise stage of the processing in the overall response. While the response RT tended to be variable (especially for trials with rather long RTs), the partition of the RT into different phases based on the eye movement status offers much cleaner data for the underlying mental process. The results suggest that the dual-state modulation is involved in the contextual cueing effect in that attentional guidance plays a major role and, to a small extent, facilitation of response selection also contributes to the overall contextual cueing effect.

Acknowledgments

This study was supported by awards from the National Key Discipline of Basic Psychology at Southwest University (NSKD11014), the Fundamental Research Funds for the Central Universities (SWU1009009), and the Natural Science and Engineer-
ing Research Council of Canada.

Author contributions: Guang Zhao and Qiang Liu contributed equally to this work.

Commercial relationships: none.

Corresponding authors: Hong Li; Hong-jin Sun.
Email: lihongwrm@vip.sina.com; sunhong@mcmaster.ca.
Addresses: Research Center of Psychological Development and Education, Liaoning Normal University, Dalian, China; Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Ontario, Canada.

References


