Spatial context and top-down strategies in visual search

ALEJANDRO LLERAS 1,* and ADRIAN VON MÜHLENEN 2

1 Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, B. C., Canada V6T 1Z4
2 Ludwig-Maximilians-University, Munich, Leopoldstrasse 13, 80802 Munich, Germany

Received 7 June 2003; revised 2 November 2003; accepted 18 November 2003

Abstract—Marvin M. Chun and Yuhong Jiang (1998) investigated the role of spatial context on visual search. They used two display conditions. In the Old Display condition, the spatial arrangement of items in the search display was kept constant throughout the experiment. In the New Display condition, the spatial arrangement of items was always novel from trial to trial. The results showed better performance with Old Displays than with New Displays. The authors proposed that repeated spatial context help guiding attention to the target location, thus they termed this effect Contextual Cueing. We present three attempts to reproduce this effect. Experiments 1 and 2 were near exact replications of experiments in Chun and Jiang’s report, where we failed to obtain Contextual Cueing. Post-experimental interviews revealed that participants used different search strategies when performing the task: an ‘active’ strategy (an active effort to find the target), or a ‘passive’ strategy (intuitive search). In Experiment 3, we manipulated task instructions to bias participants into using active or passive strategies. A robust Contextual Cueing Effect was obtained only in the passive instruction condition.

Keywords: Contextual cueing; visual search; top-down strategies; individual differences; implicit information; spatial attention.

INTRODUCTION

Our day-to-day visual environment is cluttered with a multitude of objects varying in shape, color and motion, among many other visual features. The task of finding specific objects within such a rich environment is usually guided by the features of the objects themselves but is also aided by our memory of the properties and regularities of the visual environment itself. For example, when looking for a white napkin in our kitchen, we look for white and napkin-shaped objects, but also, we most likely concentrate our attention on the part of the kitchen where we know is a napkin dispenser. Seldom will we look for napkins inside the refrigerator. This is

*To whom correspondence should be addressed. E-mail: alleras@uiuc.edu
because our knowledge about napkins and kitchens, as well as our memory of when we last saw napkins, all direct our attention to a set of locations around the kitchen with a high likelihood of holding napkins.

Myriad studies in visual cognition have focused their efforts on finding what guides visual attention through the environment under similar search-task situations (e.g. Miller, 1988; Moore and Egeth, 1998; Wolfe, 1996). The experimental task, known as visual search, consists of finding one object (the target) in an array of distracting objects (the distractors), and it is usually found that the more the target differs from the distractors (on some feature-based scale), the easier the search task becomes (e.g. Duncan and Humphreys, 1989; Treisman and Souther, 1985; Wolfe, 1994). The ensuing argument is that something like a saliency map of the items in the display helps to guide spatial attention around the display (e.g. Cave and Wolfe, 1990; Wolfe \textit{et al.}, 1989; Wolfe, 1996).

Rather than focusing on the featural relation between target and distractors (i.e. the extent to which the target resembles the distractors), Chun and Jiang (1998) investigated whether the spatial relation between target and distractors in the display can guide spatial attention. Their rationale was that the spatial arrangement of items in which a target is presented creates a spatial context for the target, much in the same way that a refrigerator, an oven and a dishwasher create a context for the location of the napkin dispenser inside a kitchen. Chun and Jiang hypothesized that if this contextual information is held constant across trials, it will guide the participants’ spatial attention more efficiently towards the location of the target. To test this hypothesis, they ran a series of 6 experiments where they manipulated the target context: targets could appear either within an ‘Old’ or a ‘New’ display configuration. In the Old display configuration trials, targets always appeared within the same spatial configuration of distractor items (there were 12 possible Old configurations), whereas in the New display configuration trials, targets always appeared within a novel spatial configuration of distractor items (there was an unlimited number of New configurations). As predicted, Chun and Jiang found reliably faster reaction times (RT) for targets appearing within the Old configurations than for targets appearing within the New configurations. They termed this RT effect Contextual Cueing. This effect has caught the attention of many researchers and has been replicated in a number of studies in the literature (e.g. Chun and Jiang, 1998, 2003; Chun and Nakayama, 2000; Endo and Takeda, 2004; Howard \textit{et al.}, in press; Hoyer and Ingolfsdottir, 2003; Jiang and Chum, 2001; Jiang and Wagner, in press; Kawahara, in press; Olson and Chum, 2002; Peterson and Kramer, 2001).

It is important to note that within Chun and Jiang’s (1998) paradigm, two things need to happen for Contextual Cueing to arise. First, participants need to learn at least a subset of the spatial configurations in which the target is consistently appearing, and second, participants need to be able to use this knowledge to guide spatial attention to the target location. If either of these two conditions is not met, Contextual Cueing cannot occur. Because we have failed to replicate the Contextual
Cueing Effect in our laboratory (see Experiments 1 and 2 below), we have been led to ask the question: how easily are these conditions met?

Regarding the first requirement for Contextual Cueing, that participants learn at least some of the repeated configurations, Chun and Jiang (1998) cleverly demonstrated that the learning of the repeated configurations occurred incidentally, that is, without the participants’ conscious effort to do so (most participants did not even notice display repetitions). Due to the nature of this incidental learning, the first requirement for Contextual Cueing seems to be met easily. However, it is difficult to independently determine whether learning has actually taken place. Indeed, Chun and Jiang also showed that memory for the repeated displays was implicit in nature (see also Chun and Jiang, 2003): despite having seen each repeated configuration over 30 times, in a post-experimental display classification task, participants were equally likely to say that an Old configuration had been presented during the experiment than they were of a random, never-before-seen, configuration.

Regarding the second requirement for Contextual Cueing (i.e. the guidance of spatial attention by memory of a previously seen configuration), Chun and Jiang’s (1998) study suggested that once the repeated configuration had been learned, guidance would occur automatically. Indeed, in every one of their experiments, Contextual Cueing occurred whenever it was predicted, and in fact in one of their experiments (Experiment 4) it even occurred with as little as one prior display exposure! Yet, if learning of the displays’ configuration is incidental and attention is so readily influenced by this knowledge, why would one fail to replicate the Contextual Cueing Effect?

Task strategies and the Contextual Cueing Effect

In this paper, we propose that participants’ search strategies can influence search performance and interfere with the guidance mechanisms involved in the Contextual Cueing Effect. Recent literature has shown that task strategy can be a strong determinant of performance, and in particular, that search strategies can influence access to implicit or pre-attentive information, as is required in Contextual Cueing. For example, Snodgrass, Shevrin and Kopka (1993) found effects of task strategy when studying unconscious perception. In their study, participants were presented with below threshold words in the form of brief 1 ms flashes preceded and followed by a bright dot. Two task strategies were used: the Pop and the Look conditions. In the Pop condition, participants were encouraged to allow words to pop into their minds following the presentation of the word, whereas in the Look condition, participants were asked to carefully examine the display in order to identify the flashed word. Evidence of unconscious perception was only found in the Pop condition, as if the Pop instructions had allowed participants to access information about the display that they could not access under Look instructions.

In a more recent study, Smilek et al. (subm.) demonstrated the effects of task strategies in a search task. In their study, participants first learned to associate
meaningful names with a few basic shapes, which were subsequently used as the stimuli in a search task. By counterbalancing shapes and meanings, Smilek et al. were able to ask one group of participants to search a display where target and distractors had similar meanings and ask a second group of participants to search an identical display in which target and distractors had very different meanings. They also manipulated search strategies. In one condition, participants were to ‘be as active as possible’ when searching for the target (Active condition), and in the second condition participants were to ‘be as receptive as possible’ when searching for the target (Passive condition). Smilek et al. found that, in the semantically dissimilar condition, RTs in the search task were strongly dependent on set size in the Active group condition; however, RTs were independent of set size in the Passive group condition. In other words, the Passive instruction manipulation turned an apparently serial search task into a pop-out task. It is as if the Passive search strategy allowed participants to use information about the displays that significantly eased their search task, information that participants in the Active search strategy did not use. See also Smilek et al. (in press) for another example of the effects of Active/Passive instruction manipulations on visual search performance.

To further investigate the lack of Contextual Cueing in our laboratory, in our Experiment 3, we used the Active/Passive instruction manipulation of Smilek et al. (subm.). If task strategies can modulate the access to implicit information, it is possible that the reason we failed to obtain Contextual Cueing in Experiments 1 and 2 is that participants were using a task strategy that blocks access to implicit information. By the same logic, one should also be able to bias participants into using a search strategy that might facilitate access to implicit information, thereby allowing a Contextual Cueing Effect to develop. To preview, our results confirmed our hypothesis: Contextual Cueing was obtained in the Passive Instruction, but not in the Active Instruction condition, which, of the two, most closely resembled the instruction conditions of Experiments 1 and 2.

EXPERIMENT 1

Experiment 1 was designed to replicate the basic Contextual Cueing Effect, by closely following the method of Chun and Jiang’s (1998) Experiment 1. In brief, no Contextual Cueing was found.

Method

Apparatus. The stimuli were presented on the 17-inch monitor of an eMac Power PC G4 (700 MHz, OS 9.2.2). The laboratory was dimly illuminated and contained 10 computers separated by partition walls. Participants viewed the monitors from a distance of 57 cm. At this distance, the total screen area subtended approximately $32^\circ \times 24^\circ$ of visual angle. Responses were recorded using the ‘z’ and ‘/’ key of the computer keyboard.
Figure 1. An example display in Experiment 1 showing randomly rotated Ls (distractors) and a T (target) pointing to the left.

Stimuli and trial sequence. The search displays always contained twelve stimuli (see Fig. 1). The stimuli were black letters, T for the target and Ls for the distractors, presented on a white background. The size of the letters was 1.4° and the thickness of the strokes 0.16° of visual angle. The Ls were randomly rotated by 0°, 90°, 180°, or 270° and the T was either rotated 90° to the left or 90° to the right. A new display was generated by placing one T and eleven Ls randomly within the cells of an imaginary 8 by 6 matrix (cell size 2.8°). Within the cell the stimuli were randomly jittered horizontally and vertically by ±0.4°.

Each trial started with a fixation cross presented for 450 ms, followed by a blank interval for 175 ms. The search display lasted for a maximum of 3000 ms, unless terminated by the participant’s response. The task was to search for an oriented T among Ls and to decide as quickly and accurately as possible whether the T was oriented to the left or to the right. After response, the display was removed and replaced by a symbol in the center of the screen providing feedback on the response (‘+’ for correct responses, ‘−’ for incorrect response, and ‘o’ for failure to respond). At the end of each block (24 trials), participants received the percentage error for this block. They were asked to try to keep the error rate below 5 percent. The inter-trial interval was 1500 ms.

Design and procedure. A one factor, within subjects design was used. The independent variable was Display type, with two levels: Old and New. Old displays had the same arrangement of distractor items in the display on every display presentation, whereas New displays had a new, random arrangement of distractor items on every display presentation. To rule out location probability effects, the target appeared equally often in each of the 24 possible locations throughout the experiment. At the beginning of the experiment, participants performed one block
with 24 randomly generated practice trials. All subsequent blocks contained the same 12 Old displays and a set of 12 randomly generated New displays, with the constraint that the first trial in every block was always a New display. There was a total of 30 blocks in the experiment, for a total of 720 experimental trials. For analysis purposes, the data from individual blocks were collapsed into 5 groups (epochs) of 6 blocks each. Dependent variables were reaction time and percentage errors.

**Instruction.** Participants received written instructions, providing a general description of the task and procedure and showing an example display. No explicit information was provided regarding strategies or the best way to solve this task. They were only instructed to respond as fast as possible, without exceeding the 5% error limit.

**Post-experimental interview.** At the end of the experiment participants where first verbally asked by the Experimenter whether they noticed that some of the displays were repeated. Then they were asked to describe how they solved the task. As we had no specific hypothesis, we used a free report approach to be as open as possible to individual differences.

**Participants.** Fourteen participants, 5 male and 9 female, took part in the experiment. Their ages ranged between 21 and 28 years. They all had normal or corrected-to-normal visual acuity and received one course credit for participating in one 45-minute session. None of the subjects were aware of the purpose of this study.

**Results**

**RT analysis.** Mean RTs were calculated for each participant and condition, excluding trials with errors or outliers (two SD below or above mean). Overall, the mean RT and standard error for Old displays was 829 ± 19 ms and the mean RT and standard error for New displays was 838 ± 22 ms, which resulted in a non-significant 9 ± 12 ms difference between Old and New display RTs.

In a second step, RTs were averaged across participants and plotted as a function of epoch, separately for old and new displays (see Fig. 2). As can be seen in Fig. 2, overall RTs became faster with increasing epoch (from 931 ± 20 ms in Epoch 1 to 773 ± 19 ms by Epoch 5). The graph also suggested that there was no difference between old and new displays across different epochs. In order to further explore the RT data, a repeated-measures analysis of variance (ANOVA) with main terms for display (old, new) and epoch (1-5) was calculated. The ANOVA confirmed the visual interpretation: there was only a significant main effects for epoch, \( F(4, 52) = 75.64, \ p < 0.001 \), but neither a main nor an interaction effect with display (both \( p > 0.32 \)). To further explore this absence of a display effect,
separate t-tests were calculated comparing old versus new displays for each epoch. The difference between old and new displays did not reach significance in any of the five epochs (all \( p > 0.19 \)).

To determine whether the absence of a significant effect was due to individual differences, the contextual cueing effect for each participant was averaged across epochs 2–4. As shown in Fig. 3, five out of fourteen participants display a negative cueing effect, one participant shows no effect, and eight participants show a positive cueing effect. This variation was not random: there was a significant within-subjects correlation of 0.74 between the mean size of the cueing effect in epochs 1 and 2 and the mean size of the cueing effect in epochs 3 and 4. In other words, most participants showing a contextual cueing effect early in the experiment, showed a contextual cueing effect in the latter part of the experiment, whereas those showing no effect (or reversed) did so also throughout the experiment.

If we removed the participants showing a negative effect from the analysis, we would get an average cueing effect of 40 ms (across epochs 2–4), which would easily reach significance, \( F(1, 8) = 14.40, p < 0.01 \). However, looking closely at the distribution of contextual cueing effects across participants we see no indication for bimodal distribution that would justify the splitting of subjects into two groups.
Table 1.
Mean percentage errors in Experiment 1 for old/new displays, separate for each epoch

<table>
<thead>
<tr>
<th>Display</th>
<th>Active Old</th>
<th>Old</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epoch 1</td>
<td>3.2</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>Epoch 2</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Epoch 3</td>
<td>3.2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Epoch 4</td>
<td>2.2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Epoch 5</td>
<td>2.3</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.6</td>
<td>3.0</td>
<td></td>
</tr>
</tbody>
</table>

There is only a weak correlation between overall RT and the contextual cueing effect ($r = 0.23$). This rules out the possibility that only participants with slow RTs showed a contextual cueing effect.

Error analysis. Overall percentage error for each participant in each condition was averaged across participants (see Table 1). The overall mean error rate was relatively low (2.8 ± 0.8%), indicating that participants had no trouble keeping the error rate below the 5% limit. Errors were arcsine-transformed and analyzed by a two-way repeated-measure ANOVA with main terms for display (old, new) and epoch (1-5). None of the main effects or interaction effects reached significance (all $p > 0.14$). There was also no correlation between accuracy and the size of the contextual cueing effect ($r = 0.03$). It seems therefore unlikely that our RTs are confounded by differential speed–accuracy trade-offs.

Post-experimental interview. Of the fourteen participants, three reported to have noticed that some displays were repeated throughout the experiment. However, they estimated that on average less than 20% of the displays were repeated. There seems to be little evidence that conscious awareness of the repetition of some displays has had an influence on the contextual cueing effect, as two of them showed a positive cueing effect and one showed a negative cueing effect. Regarding task strategies, participants’ subjective reports were quite varied. Some participants tried very systematic approaches, others used more ‘intuitive’ approaches to the task, and a third group of subjects were reluctant to name a specific strategy. We tried to order the various reports into three groups: Active search strategy (3), Passive search strategy (4), no strategy/report (7). Only one of the participants using an active search strategy showed a contextual cueing effect, whereas three participants using a passive search strategy showed a contextual cueing effect. This posthoc result is consistent with the results we later obtained in Experiment 3.
Discussion

Experiment 1 failed to find a robust contextual cueing effect. There was a small trend on the right direction on epochs 2 to 5, but the effects were in the range of 10 to 20 ms and not significant. In comparison, in their Experiment 1, Chun and Jiang (1998) had reliable effects in the range of 60 to 80 ms. It must be noted, however, that our methods were slightly different from those of Chun and Jiang in two minor ways. The most noticeable difference was the color of the stimuli: in Chun and Jiang’s study, stimuli were green, red, blue or yellow, whereas in our displays all stimuli were black. The second difference between our studies was the background of the displays, which was textured in Chun and Jiang’s experiment and white in ours. However, in related experiments Chun and colleagues have shown that these minor differences are not critical in obtaining Contextual Cueing (e.g. Chun and Jiang, 1998, 2003; Olson and Chun, 2002). Thus, we feel confident that these minor differences in the display are not to blame for the lack of a Contextual Cueing Effect.

Another possible reason for the lack of contextual cueing is that participants in our experiment might have found the task relatively easier than the participants in Chun and Jiang’s study. In our experiment, the average RTs were in the range of 770 to 930 ms, while in Chun and Jiang’s Experiment 1 RTs were slightly longer, in the range of 800 to 1080. So, it is possible, yet unlikely, that our participants might have been more efficient in the search task, not leaving enough room for the contextual cueing effect to play a role. This possibility was taken into consideration in the design of Experiments 2 and 3, where the search task was made more difficult.

A third (and in our opinion more likely) reason that may account for our missing Contextual Cueing Effect is that some participants may have been using a specific search strategy that produced no contextual cueing effect, or even a negative contextual cueing effect. The post-experimental interviews suggested that participants using a passive search strategy may show more contextual cueing than participants using an active search strategy. Experiment 3 was designed to directly test this hypothesis.

EXPERIMENT 2

Experiment 2 was identical to Experiment 1, except that the search task became more difficult. This was done by shifting the vertical line of the L-shaped stimuli by 0.3° to the right in order to make it more similar to the T-shaped targets. To allow time for the participants to find the target, the trial cut off was increased from 3000 to 4500 ms. Fifteen participants participated in the experiment.

The mean RT for Old Displays was 2003 ± 67 ms and the mean RT for New Displays was 2007 ± 67 ms (a 4 ± 38 ms difference between the two conditions). Despite the more difficult task, the overall result showed no significant Contextual Cueing effect. In addition, the epoch analysis showed no Contextual Cueing effect at any of the epochs either (average 7 ms difference across epochs 2–5).
overall error rate was higher than in Experiment 1, but did not differ across display condition (mean error rates of 12.1 ± 2.3% for Old Displays and 12.8 ± 2.1% for New Displays).

Both RTs and accuracy measures reflected the higher difficulty of the search task in Experiment 2: RTs in Experiment 2 were over a second longer than RTs in Experiment 1, and participants in Experiment 2 were 10% less accurate too. Even though these long response times should have been sufficient to allow for a Contextual Cueing Effect to emerge, none was obtained. Thus, we feel confident that target discriminability was not responsible for the absence of Contextual Cueing in Experiment 1. Note that this result is consistent with the observations of Olson and Chun (2002) that spatial features, rather than surface features, determine Contextual Cueing.

EXPERIMENT 3

The goal of Experiment 3 was to evaluate the effect of search strategies on the Contextual Cueing Effect. Two sets of task instructions were used in order to bias participants into using either an ‘active’ or a ‘passive’ search strategy while doing the task (see Smilek et al., subm.). Overall, the instruction manipulation was successful and a robust Contextual Cueing Effect was found for the ‘passive’ instruction group, while no effect was found for the ‘active’ instruction group.

Method

The method in Experiment 3 was almost identical to Experiment 1 and 2, except for the following: participants received different instructions (see below), and the distractor stimuli were slightly changed. As in Experiment 2, the vertical line of the L-shaped stimuli was somewhat shifted to the right, but this time only by 0.1° (instead of 0.3°). This change increased the overall difficulty of the search task a little bit over Experiment 1, but not as much as in Experiment 2. The maximum trial duration was 4500 ms. Moreover, only 24 blocks (instead of 30 blocks) were presented, in order to keep the overall length of the experiment below one hour. In addition, at the end of the experiment, participants received a short questionnaire, with the following questions: (a) did you find the instructions helpful, (b) did you notice that some of the displays were repeated throughout the experiment, and (c) if so, what proportion of the displays (%) would you say were repeated?

Instructions. Participants were randomly assigned to one of two conditions: ‘active search strategy instructions’ or ‘passive search strategy instructions’. While administering the instructions, experimenters showed participants a sample display with a general description of the task and the procedure. Furthermore, participants in the active strategy instruction group received the following set of instructions, which were adapted from Smilek et al. (subm.):
“The best strategy for this task, and the one that we want you to use in this study, is to be as active as possible and to ‘search’ for the item as you look at the screen. The idea is to deliberately direct your attention to determine your response. Sometimes people find it difficult or strange to ‘direct their attention’, but we would like you to try your best. Try to respond as quickly and accurately as you can while using this strategy. Remember, it is very critical for this experiment that you actively search for the unique item”.

Participants in the passive strategy instruction group received the following set of instructions:

“The best strategy for this task, and the one that we want you to use in this study, is to be as receptive as possible and let the unique item ‘pop’ into your mind as you look at the screen. The idea is to let the display and your intuition determine your response. Sometimes people find it difficult or strange to tune into their ‘gut feelings’, but we would like you to try your best. Try to respond as quickly and accurately as you can while using this strategy. Remember, it is very critical for this experiment that you let the unique item just ‘pop’ into your mind’.

Participants. Forty new participants, 13 male and 27 female, took part in the experiment. Their ages ranged between 18 and 26 years. They all had normal or corrected-to-normal visual acuity and received one course credit for participating in one 45-minute session. None of the subjects was aware of the purpose of this study.

Results

RT analysis. Mean RTs were again calculated for each participant for each display condition and instruction group. Overall, in the active-search group, the mean RT for Old displays was 1216 ± 54 ms and the mean RT for New displays was 1237 ± 55 ms, which resulted in a non-significant 21 ± 19 ms difference between Old and New display RTs. On the other hand, in the passive-search group, the mean RT for Old displays was 1196 ± 53 ms and the mean RT for New displays was 1302 ± 60 ms, which resulted in a significant 106 ± 28 ms Contextual Cueing Effect.

As in Experiment 1, the data was also averaged across participants and plotted as a function of epoch (see Fig. 4), separately for the active-search (top graph) and passive-search (bottom graph) instruction. As can be seen in both graphs, overall RTs decreased with increasing epoch, showing a performance improvement close to 200 ms over the length of the experiment. Consistent with the pattern observed in the overall means, RTs were faster with old than with new displays, but only in the passive and not in the active instruction group. This interpretation of the figure was confirmed by a three-way mixed-design ANOVA with the between-subject factor Instruction condition (active and passive), and the within-subject factor display (old, new) and epoch (1-4). There were significant main effects
Figure 4. Mean correct reaction times (RTs) and associated standard errors in milliseconds as a function of epoch in Experiment 3 and instruction group. (A) Left panel: Active instruction group. (B) Right panel: Passive instruction group. Circles represent the old display condition and triangles the new display condition.

for epoch, $F(3, 114) = 142.02, p < 0.001$, and for display, $F(1, 38) = 16.33, p < 0.001$. Furthermore, the display x search strategy interaction was significant, $F(1, 38) = 16.33, p < 0.001$. None of the other main or interaction effects were significant. Two follow-up ANOVAs were conducted separately for the active and for the passive search strategies to further explore the display x search strategy interaction. Both ANOVAs revealed a significant main effect for epoch; however, only the passive search strategy condition showed a significant effect for the display type, $F(1, 19) = 20.01, p < 0.001$, and not the active search strategy condition, $F(1, 19) = 1.05, p = 0.32$, confirming that a Contextual Cueing Effect was observed only in the passive search strategy condition.

Participants in the passive condition showed a significant display type effect as early as in the first epoch, $t(19) = 3.70, p < 0.01$. To further explore this effect, we looked separately at the data from each of the 6 blocks of the first epoch. Note that there were only 12 trials for each condition. The block-by-block analysis showed that although there seemed to be a positive cueing effect in the first block of the passive condition and a negative cueing effect in the first block of the active condition, neither of these differences was significant. More remarkably, however, there was a significant or marginally significant positive Contextual Cueing effect in blocks 2 through 6 of the passive condition (all $p < 0.1$), whereas there was no significant cueing (either positive or negative) in the corresponding blocks of the active condition.

Last, to further illustrate the effectiveness of the instruction manipulation, we analyzed individual subject data (see Fig. 5). Only 35% of subjects in the Active instruction group showed a positive contextual cueing effect, whereas 80% of subjects in the Passive instruction group showed positive Contextual Cueing. To verify whether participants kept a consistent search strategy throughout the experiment, we calculated a within-subject correlation between the mean size of the cueing effect in epochs 1 and 2 and the mean size of the cueing effect in epochs 3 and 4. For the active search group, the correlation was 0.81, while, for the passive search group, the correlation was 0.67. Both correlations are substantial and highly
Figure 5. Contextual cueing effect (in milliseconds) in Experiment 3, separate for each participant. (A) Top panel: Active instruction group. (B) Bottom panel: Passive instruction group.

significant, so we feel confident that participants used a consistent search strategy (active or passive) throughout the experiment.

In summary, the overall pattern of results confirmed that our instruction manipulation strongly influenced the participants’ performance in the search task. Furthermore, at an individual level, the results showed that Contextual Cueing is more than twice as likely to arise with a Passive rather than with an Active search strategy.

Error analysis. The overall error rate was very low in both conditions: 1.8 ± 0.4% in the active condition and 2.4 ± 0.5% in the passive condition (see Table 2). Errors were arcsine-transformed and analyzed in a three-way mixed-design analysis of variance (ANOVA) with the between-subject variable search strategy (active and passive), and the within-subject display (old, new) and epoch (1-4). None of the main effects or interaction effects reached significance (all $p > 0.11$). It is therefore unlikely that our RTs in Experiment 2 are confounded by differential speed-accuracy trade-offs.
Table 2.
Mean percentage errors for active/passive search strategies and old/new displays, separate for each epoch

<table>
<thead>
<tr>
<th>Display</th>
<th>Search strategy</th>
<th>Active</th>
<th>Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Old</td>
<td>New</td>
</tr>
<tr>
<td>Epoch 1</td>
<td></td>
<td>1.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Epoch 2</td>
<td></td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Epoch 3</td>
<td></td>
<td>1.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Epoch 4</td>
<td></td>
<td>1.7</td>
<td>1.8</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>1.8</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Post-experimental questionnaire. Only two participants (in the Passive group) out of forty found the instructions not helpful. Five participants in the active and 6 participants in the passive search condition reported to have noticed the repetition of some displays. Of all these participants, only two found the correct percentage of repetition (50%). Overall, the questionnaire did not reveal any systematic differences between the active and passive search condition participants.

Discussion

We believe that the instruction manipulation in Experiment 3 successfully biased participants into adopting different search strategies. While most participants in the Passive instruction group used a search strategy that yielded a Contextual Cueing Effect, most participants in the Active instruction group used a search strategy that yielded no Contextual Cueing.

The results of Experiment 3 confirmed our interpretation of Experiment 1: that no Contextual Cueing Effect was found there because some of the participants were using a search strategy that did not allow for such an effect to emerge. Indeed, in Experiment 1, 43% of the subjects that were able to retrospectively describe their search strategy (3 out of 7) described a strategy more consistent with the Active search strategy of Experiment 3. Insofar as we can take the few successful self-reports of Experiment 1 as an indicator of the overall proportion of participants using an Active-type search strategy, we can infer that about half of the subjects in Experiment 1 did use an Active-type search strategy, and as the results of Experiment 2 have shown, an Active-type search strategy does not easily yield Contextual Cueing. Coincidentally, also 43% of subjects in Experiment 1 (6 out of 14) failed to show Contextual Cueing.
GENERAL DISCUSSION

To-date, there have been many reports of Contextual Cueing in the literature (Chun and Jiang, 1998, 2003; Chun and Nakayama, 2000; Howard et al., in press; Hoyer and Ingolfsdottir, 2003; Jiang and Chun, 2001; Kawahara, in press; Manns and Squire, 2002; Olson and Chun, 2002; Peterson and Kramer, 2001), attesting to the robustness and importance of this phenomenon. Yet, when we tried to replicate it in our laboratory, we failed to obtain it (see Experiments 1 and 2). This led us to ask the following questions: What are the task requirements to obtain Contextual Cueing? And why were they not met in our Experiments 1 and 2, when they seemed to be met under quasi-identical experimental conditions in other laboratories?

As first mentioned in the introduction, at least two separate things need to happen for Contextual Cueing to arise. First, some learning of the spatial configuration of items in the Old displays must occur. Whether participants need to learn the entire display configuration or just the local context around the target is debatable (see Olson and Chun, 2002; Jiang and Wagner, in press); however, some learning of the repeated display configurations must occur or otherwise the visual system will not recognize the repeated displays when they re-appear later in the experiment. Second, the learned information about the displays must be able to guide spatial attention to the target location, or in other words, spatial attention has to be susceptible to the influence of the implicitly learned information about spatial context in repeated displays. If, for any given reason, spatial attention either does not have access to or is not susceptible to the learned spatial configuration information, there can be no contextual guidance to the target location.

Experiment 3 provided us with a possible answer to our question of why we did not get Contextual Cueing in Experiments 1 and 2: like participants in the Active group of Experiment 3, participants in Experiments 1 and 2 might have used a search strategy that did not allow for the implicit information about spatial context to guide attention. However, a lack of Contextual Cueing might also have arisen because Active-like search strategies prevented implicit learning from occurring. Although a possibility, a lack of learning seems to be a rather unlikely account for these results, particularly given the documented ease of the learning in question. Not only is the learning involved in Contextual Cueing incidental (i.e. effortless for the participants), but also the influence of this knowledge on performance can occur quite soon after the initial learning takes place: Chun and Jiang’s (1998) results, as well as the results of the Passive condition of Experiment 3 both have shown that sometimes a single glance at a search display (among 24 other displays) is sufficient for the learning of the spatial configurations to take place, such that this learned information can have an effect on performance from the very first time Old displays are repeated! In other words, participants encoded and learned the spatial layout of items in the displays even before any of the displays were repeated, that is, even before the spatial context became a relevant feature in the search task.

It is possible that implicit learning of spatial configurations might have taken place, but in a fragmented (rather than wholesome) manner. That is, because Active-
like strategies encourage the rapid deployment of attention from item to item across
the display, spatial context might be learned in fragments, or dislocated islets of
items, with no direct link to the target location. However, recent data by Olson and
Chun (2002) showed that, even if this were the case, the lack of global contextual
information would not suffice to eliminate Contextual Cueing. Indeed, Olson and
Chun found that local contextual information alone could produce a Contextual
Cueing Effect. Thus, even if Active-type strategies do dislocate the display into
smaller islets of items, the contextual information in and immediately around the
target islet should have been sufficient to yield Contextual Cueing.

Lastly, it is important to note that the learning of spatial configurations is not only
effortless but also very robust: Contextual Cueing is still present even one week after
the initial exposure to the test displays (Chun and Jiang, 2003). Thus, the learning
mechanism involved in Contextual Cueing is one that can very easily extract spatial
information from the displays, both globally and locally, and quickly cement it into
a very robust representation that can be readily accessible, even many days after
the initial testing. We therefore believe that it is unlikely that this kind of learning
mechanism would have faltered in our Experiments 1 and 2 and only in one of the
two conditions of Experiment 3.

If we believe that implicit learning of spatial configurations did take place
in Experiment 3, why would an Active search strategy prevent access to this
information, whereas a Passive search strategy fosters it? This might come to
happen because under an Active strategy, participants might feel compelled to
actively control and direct the focus of their attention during each trial, thereby
blocking any guidance from implicitly learned information. In other words, because
participants are unaware of having gathered task-relevant contextual information
about the location of the target in the search displays, they fail to use it. In
contrast, Passive-like search strategies might allow attentional guidance from
implicit information because under a Passive strategy, participants might be more
willing to surrender active control of their focus of attention, thereby allowing
implicit knowledge (maybe in the form of hunches) to guide them around the
display. Note that we do not mean to say that only implicit knowledge can guide
spatial attention under Passive strategies. Rather, we believe that a Passive strategy
simply allows implicit knowledge to be one of the many guiding mechanisms of
spatial attention.

In summary, we propose that the absence of Contextual Cueing in our experiments
probably does not reflect a lack of implicit learning, but rather a lack of attentional
guidance: participants most likely did learn contextual information about the
displays, but this information failed to have an influence on their performance
because of the search strategy they used. At this point however, we cannot
provide strong empirical support for this conclusion, yet we believe that a lack
of attentional guidance is the hypothesis most consistent with previous studies
on Contextual Cueing (Chun and Jiang, 1998, 2003; Olson and Chun, 2002) as
well as with previous research on how task strategies differentially affect access
to implicit information (Smilek et al., subm.; Snodgrass et al., 1993). Although other researchers have extensively studied Contextual Cueing, they have all failed to explicitly discuss the link between the implicit learned contextual information and spatial attention, assuming that implicit learning meant automatic attentional guidance and that a lack of attentional guidance implied an absence of implicit learning. The new findings in this manuscript imply that, unlike previously assumed, implicitly learned contextual information may not automatically guide spatial attention, a fact that should be taken into consideration in future research on Contextual Cueing.

CONCLUSION

The present study illustrates how instruction manipulations can influence performance in visual search tasks in general and in Contextual Cueing, in particular. While a passive instruction set allows for substantial Contextual Cueing effects to arise, an active instruction set can eliminate this effect. Thus, insofar as precise instruction sets can accurately determine participants search strategies, a lack of precise and consistent information in instruction sets allows participants to arbitrarily select one of many search strategies. Such freedom of choice can be costly as participants might end up choosing a search strategy that conflicts with the main goal of the investigation, just as active search strategies conflicted with our investigation of Contextual Cueing. We hope this study will provide investigators with an instruction set that will allow them to more consistently obtain and more precisely study Contextual Cueing. Lastly, we believe this study will open the door to further investigations into the links between implicit information and guidance of spatial attention, as such links might turn out to be more complex than originally thought.

Acknowledgements

This research was supported by an NSF fellowship to Alejandro Lleras and a Deutsche Forschungsgemeinschaft grant FOR 480 and a Swiss National Science Foundation fellowship to A. von Mühlenen. We would like to thank Dan Smilek for helping us with formulating the active/passive instructions.

REFERENCES


