Use of spatial context is restricted by relative position in implicit learning

NOBUTAKA ENDO and YUJI TAKEDA

National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

For inefficient search, target detection is faster for repeated than for regenerated layouts. This effect, called *contextual cuing*, was assumed to arise from implicit learning of local spatial relationships between targets and distractors. However, a more global influence from distractors far from the target has not been tested. In this study, the search field was divided into upper and lower halves containing a repeated and a regenerated configuration set, respectively. The positions of the two sets were or were not exchanged, meaning that their relative as well as their absolute positions were the same or different (Experiment 1). In Experiment 2, the repeated set appeared alone in either the same or the other half of the screen (same or different absolute position). The contextual cuing effect remained when only absolute position was changed, but not when both absolute and relative positions were changed. These results suggest that contextual cuing depends on relative positional information.

Visual context, such as the spatial configuration of objects, is a critical factor for the guidance of attention to a target location (Biederman, 1972; Biederman, Mez-zanotte, & Rabinowitz, 1982). For example, Biederman et al. (1982) showed that a target placed in an irregular location was harder to detect than a target placed in its typical location. Recently, Chun and his colleagues have developed a paradigm using a visual search task to examine how the visual context of the search display affects search performance. In a series of studies, they demonstrated that the association between a target location and the distractor configuration is encoded when the same display is presented repeatedly, and that the configuration can be used as a cue for the target location in latter blocks, even if observers are unaware that the configurations are repeated (Chun, 2000; Chun & Jiang, 1998, 2003; Chun & Nakayama, 2000; Olson & Chun, 2002). This facilitatory effect is termed *contextual cuing*.

In the paradigm used by Chun and Jiang (1998), participants were asked to search for a target (e.g., *T*) among distractors (e.g., *Ls*). There were two conditions: the *old* and *new* conditions. In the *old* condition, the search array presented in the first block was repeatedly presented throughout the latter blocks (i.e., the locations of the target and distractors were fixed). In the *new* condition, only

the target location in a given search array was fixed across the blocks, and the distractor locations were randomly changed in every block. The results showed that reaction times (RTs) became gradually faster in the *old* condition than they were in the *new* condition. Participants did not notice the repetition of the search arrays in the *old* condition, indicating that the visual context (i.e., the association between the target location and the distractor configuration) was implicitly learned and was used to search for a target.

How is visual context represented? Recently, some studies have investigated the nature of the representation of spatial configuration. Jiang and Wagner (2004, Experiment 2) demonstrated that contextual cuing could be observed even when the search display was rescaled from $18^\circ \times 12^\circ$ of visual angle in the training phase to $22.5^\circ \times 15^\circ$ (or to $14.4^\circ \times 9.6^\circ$) of visual angle in the test phase, suggesting that the visual context is encoded on the basis of the relative positions of the search items rather than of the absolute position of each item in the CRT display. This account is also supported by the results of a study using 3-D search arrays. Chua and Chun (2003) investigated whether contextual cuing was affected by increasing the angle of rotation in depth (0° , 15° , 30° , and 45°) away from the training views. They reported that contextual cuing for the 3-D configuration was observed in the training phase, but that the learning effect decreased and disappeared with increases in the rotation angle. This finding indicates that the representation of spatial configuration is viewpoint dependent, so that contextual cuing is weakened by warping the relative positions of the search items in the 2-D image.

In most studies of spatial contextual cuing, including the studies described above, the whole configuration of the search array was repeated throughout the experimental blocks. However, it has been reported that the repetition

This study was supported by a research fellowship to N.E. from the Japan Society for the Promotion of Science for Young Scientists. We thank Cathleen Moore, Thomas Sanocki, Ingrid Olson, and an anonymous reviewer for useful comments on earlier versions of this article. We are also grateful to Steven Phillips for English phraseology. Correspondence relating to this article may be sent to N. Endo, Cognition and Action Research Group, Institute for Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology, Tsukuba Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki, 305-8566 Japan (e-mail: n-endou@aist.go.jp).

of the whole display is unnecessary for contextual cuing. Olson and Chun (2002) showed contextual cuing when only the half of the display configuration that included the target was repeated, and the other half was randomly regenerated. Furthermore, the magnitude of the contextual cuing effect with only half of the configuration repeated was comparable to the magnitude when the whole configuration was repeated. These results indicate that visual context (i.e., repeated configuration) can be extracted from noise (i.e., regenerated configuration) and can be used as a cue for the target location.

The purpose of the present study, similar to that of Olson and Chun (2002), was to examine how visual context is represented when only half of a configuration is repeated. As was shown by Jiang and Wagner (2004), contextual cuing relies on encoding the positions of items within the repeated region of the display. However, it is not clear whether such cuing also depends on encoding the position of the noise region (i.e., the region containing regenerated items) relative to the repeated region. If the spatial relationship between context and noise is encoded, contextual cuing will be restricted by the relative position of the noise region; otherwise, the relative position of the noise should have no effect. In the present study, we examined the robustness of contextual cuing under conditions in which the relative and absolute positions of the visual context varied. That is, repeated items (the visual context) maintained the same position relative to each other but changed position relative to the noise set (the relative position factor) or only changed their location on the CRT display (the absolute position factor). In Experiment 1, we exchanged the positions of the repeated configuration and the noise configuration. In Experiment 2, we varied only the absolute position of the visual context by shifting the position of the repeated search array.

EXPERIMENT 1

In Experiment 1, we investigated whether the use of the learned configuration was restricted by its position. In the training phase, participants performed the search task with and without repeated context in the *old* and *new* conditions. In the *old* condition, either the upper or the lower half of the display configuration was repeated and the remaining half of the configuration was randomly regenerated, in a manner similar to Olson and Chun's (2002) Experiment 2. In the *new* condition, the entire display configuration was randomly regenerated for each trial. The test phase consisted of three types of repetition condition: *old*, shift, and new-old. The *old* condition was the same as that in the training phase. In the shift condition, the learned configuration from the training phase was repeated, but the positions of the learned configuration and the noise configuration were exchanged—that is, both the relative and absolute positions of the visual context changed. In the new-old condition, a new configuration was generated at the beginning of the test phase, and the half of the configuration containing the target was repeated for the rest of the phase. Although there are many

ways to change relative position, we tested the simplest case in order to clarify the basic effect of position change on contextual cuing.

If contextual cuing occurs regardless of the position of the learned configuration during the training phase, performance in the shift condition should be comparable to that in the *old* condition. Alternatively, if contextual cuing does depend on the position of the learned configuration, performance in the shift condition should be similar to that in the new-old condition.

Method

Participants. Twenty paid volunteers (20–28 years old) participated in Experiment 1. All participants had self-reported normal or corrected-to-normal vision.

Stimuli and Apparatus. Each search display contained 16 items, 1 target and 15 distractors, presented at randomly selected locations within an invisible 22×16 matrix ($31.2^\circ \times 22.7^\circ$). Each item subtended $1.15^\circ \times 1.15^\circ$. The target was a rotated *T* (90° or 270°), and the distractors were rotated *Ls* (0° , 90° , 180° , and 270°). To make the search task difficult, the distractor *Ls* had a small offset (0.14°) at the line junctions. The search display was split into equal upper and lower halves by an invisible horizontal line. Each half contained eight items (see Figure 1). The search arrays were displayed on a 17-in. CRT monitor with $1,024 \times 768$ resolution connected to a computer. The participants responded using a mouse button, and RTs and responses were recorded on the computer. A program written in MATLAB controlled the experimental schedule, using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). All items were presented in white on a gray background in a semidark room. The participants observed the search array at a distance of 57 cm.

Design. The participants performed 24 blocks of trials. The first 12 blocks were the training phase, and the last 12 were the test phase. In the training phase, each block consisted of 48 trials, divided into two repetition conditions: *old* (32 trials) and *new* (16 trials). The target location in each trial was selected randomly at the beginning of the experiment, and these target locations were repeated across blocks. In the *old* condition, either the upper or the lower half of the configuration (the one that included the target location) was repeated across the training phase, while the remaining half of the configuration was regenerated. In the *new* condition, only the target location was repeated; the locations of all distractors were randomly regenerated in each trial.

In the test phase, each participant was tested in three repetition conditions: *old*, shift, and new-old. In all conditions, the half of the screen containing the target location was repeated across the test phase, while the other half was randomly regenerated. The 32 configurations from the *old* trials in the training phase were equally assigned to either the *old* or the shift condition in the test phase on the basis of mean RTs in training, to correct for any bias in mean RTs between the *old* and shift conditions. The *old* condition (16 trials) was identical to that in the training phase. In the shift condition (16 trials), the location of the repeated half swapped sides so that, for example, a repeated configuration in the upper half of the display in the training phase (including the target) occupied the lower half of the display in the test phase. In the new-old condition (16 trials), the half containing the target location was regenerated at the beginning of the test phase and was repeated for all remaining blocks. This condition accounted for learning during the test phase, and so provided a baseline to compare with the effects of transfer from the training phase in the *old* and shift conditions. A sample of these repetition conditions is shown in Figure 1.

Procedure. Each trial started with a white fixation cross at the center of the display, which remained until the end of the trial. After 1,000 msec, a search array was presented until the participant re-

Figure 1. Experimental manipulations of spatial configuration for each repetition condition across the training and test phases. Random configurations that did not include the target location were not presented in Experiment 2, but that half of the display was instead left blank.

sponded. The participant was required to search for a target and to respond with its orientation (left or right) as quickly and accurately as possible by pressing the mouse button. After a response, the search array and the fixation cross were cleared, and the next trial followed 2,000 msec later. If the participant's response was incorrect, a beep sounded as feedback. At the end of the visual search task, a recognition test was conducted for the 16 configurations used in the *old* condition at test and 16 newly generated configurations in order to test whether the participants had memorized the configurations. Before the experimental session, each participant practiced for 48 trials. The target locations and the distractor configurations in the practice session were different from those used in the experimental session.

Results and Discussion

To increase statistical power, the 24 blocks were divided into eight epochs of 3 blocks each. Incorrect or outlier trials were excluded from the analysis. Outliers were defined as trials with RTs more than three standard deviations from the mean in each repetition condition in each epoch. This procedure was also used in Experiment 2. The data of 1 participant were excluded from the analysis because more than 10% of the trials were errors or outliers. The mean rate of excluded trials across epochs was 2.32%. Mean correct RTs and mean error rates as functions of epoch and condition are shown in Figure 2 and Table 1, respectively.¹ For the recognition test, the mean hit and false alarm rates were 59.19% and 47.06%, respectively.²

RTs in the training phase were subjected to a two-way ANOVA with epoch (1–4) and condition (*old* and *new*) as factors. The main effect of epoch was significant [$F(3,54) = 7.07, p < .001$]. The main effect of condition was also significant [$F(1,18) = 14.92, p < .005$], indicating that RTs were faster in the *old* than in the *new* condition. The interaction between epoch and condition was marginally significant [$F(3,54) = 2.60, p = .061$]. A Newman–Keuls test indicated that RTs in Epoch 4 were faster in the *old* than in the *new* condition ($p < .01$).

The transfer effect in the test phase was also assessed. RTs in the test phase were subjected to a two-way ANOVA with epoch (5–8) and condition (*old*, *shift*, and *new-old*) as the factors. The main effect of epoch was significant [$F(3,54) = 5.75, p < .005$], as was the main effect of condition [$F(2,36) = 7.05, p < .005$]. A Newman–Keuls test indicated that RTs were faster in the *old* condition than in the *shift* and *new-old* conditions (all $ps < .05$). However, the difference in RTs between the *shift* and *new-old* conditions was not significant ($p = .17$), nor was the interaction between epoch and repetition.

In the training phase, RTs were shorter in the *old* than in the *new* condition. This result is consistent with the results of Olson and Chun's (2002) study, suggesting that a localized configuration set is extracted from the entire configuration and is sufficient for contextual cuing. In contrast, the learning effect was completely abolished

Figure 2. Mean reaction times as a function of epoch for each repetition condition, illustrating the learning function in Experiment 1. Error bars represent one standard error.

by shifting the position of the learned configuration on the display between the training and test phases. Thus, it is likely that the use of the learned configuration is restricted by its position relative to the noise configuration. However, the manipulation in Experiment 1 confounded relative and absolute position changes, so it is also possible that contextual cuing was affected by the absolute position change. In Experiment 2, we attempted to determine which of these position changes was affecting the occurrence of contextual cuing.

EXPERIMENT 2

The results of Experiment 1 indicate that the occurrence of contextual cuing was affected by exchanging the

positions of the learned configuration and the noise configuration. However, it is unclear whether the relative or the absolute position change was what restricted the use of the learned configuration in visual search.

In Experiment 2, we manipulated only the absolute position change and examined the occurrence of contextual cuing in *old* and *shift* conditions without a noise configuration.

Method

Participants, Stimuli, and Apparatus. Twenty-five paid volunteers (19–28 years old) participated in Experiment 2. All participants had self-reported normal or corrected-to-normal vision. The apparatus and stimuli were identical to those in Experiment 1, except that only the repeated half of the display containing the target location was presented (see Figure 1). Thus, the number of search items was reduced from 16 to 8. The search items were presented in either the upper or the lower half of the display, and the other half of the display was always empty.

Design and Procedure. The experimental design was identical to that in Experiment 1. In the training phase, the participants were tested in two repetition conditions: *old* and *new*. In the *old* condition, the target location and the distractor configuration were repeated across the training phase without any noise configuration, and in the *new* condition, the target location was repeated, but the distractor configuration was regenerated for every block.

In the test phase, the *old* condition was identical to that in the training phase. In the *new-old* condition, the distractor configuration was regenerated at the beginning of the test phase and was repeated across all blocks of trials. Finally, in the *shift* condition, the target location and the distractor configuration were shifted to the opposite half of the display. Thus, in this condition, only the absolute positions of the target and the distractors changed between the training and test phases.

Results and Discussion

The data of 1 participant were excluded from the analysis because more than 10% of the trials were errors or outliers. The mean rate of excluded trials across epochs was 2.50%. Mean correct RTs and mean error rates as functions of epoch and condition are shown in Figure 3 and Table 1, respectively.³ For the recognition test, both the mean hit and false alarm rates were 52.60%, indicating that contextual cuing involved implicit memory for the learned configurations.

RTs in the training phase were subjected to a two-way ANOVA with epoch (1–4) and condition (*old* and *new*) as the factors. The main effect of epoch was significant [$F(3,69) = 26.38, p < .001$]. The main effect of condition was also significant [$F(1,23) = 7.66, p < .05$], indicating that RTs were faster in the *old* than in the *new* condition. The interaction was marginally significant [$F(3,69) = 2.24, p = .091$]. A Newman–Keuls test indicated that RTs in Epochs 3 and 4 were faster in the *old* than in the *new* condition (all $ps < .05$).

RTs in the test phase were also subjected to a two-way ANOVA with epoch (5–8) and condition (*old*, *shift*, and *new-old*) as the factors. The main effect of epoch was significant [$F(3,69) = 7.79, p < .001$], as was the main effect of condition [$F(2,46) = 3.59, p < .05$]. A Newman–Keuls test indicated that RTs in the *new-old* condition were slower than in the *shift* condition ($p < .05$) and tended to be slower than in the *old* condition ($p = .07$).

Table 1
Mean Error Rates (%) as a Function of
Epoch and Repetition Condition

Condition	Training Phase				Test Phase			
	1	2	3	4	5	6	7	8
Experiment 1								
New	0.99	0.77	0.88	0.33				
New-old					0.99	0.55	0.55	0.44
Old	0.99	0.71	0.60	0.66	0.44	0.55	0.66	0.44
Shift					0.55	0.66	0.66	0.88
Experiment 2								
New	1.13	1.22	0.95	1.39				
New-old					1.04	1.22	1.13	1.04
Old	1.61	0.87	1.13	0.95	1.48	1.04	1.04	1.39
Shift					1.82	1.48	1.04	1.04

Figure 3. Mean reaction times as a function of epoch for each repetition condition, illustrating the learning function in Experiment 2. Error bars represent one standard error. Note that the number of search items was 8 in Experiment 2, where it had been 16 in Experiment 1. Thus, mean reaction times in Experiment 2 were shorter than those in Experiment 1.

However, the difference in RTs between the *old* and shift conditions was not significant ($p = .45$), nor was the interaction between epoch and condition.

In contrast to the results of Experiment 1, Experiment 2 demonstrated that the learning effect of the repeated configurations persisted even when the absolute positions were changed between the training and test phases. These results suggest that the occurrence of spatial contextual cuing is not affected by the absolute position of the learned configuration.

GENERAL DISCUSSION

The purpose of this study was to determine whether the use of learned context is restricted by the position it was displayed in during the learning phase. In Experiment 1, spatial contextual cuing was observed even when half of the display was a random configuration in the training phase, but the learning effect was not observed when the relative positions of the learned configuration and a noise configuration were exchanged in the test phase. In contrast, in Experiment 2, spatial contextual cuing was observed regardless of the absolute position of the learned configuration in the CRT display. These results suggest that the occurrence of contextual cuing depends on the relative positions of a learned configuration and a noise configuration.

One may question whether the learned configuration always guided observers' attention to the location where the

target had been presented in the training phase. Thus, in the shift condition in Experiment 1, the contextual cuing effect disappeared when the location of a learned context was shifted in the test phase. However, the results of our pilot study showed that visual search was not facilitated even when the target was presented at the original location in the shift condition. Overall, the results of the present study indicate that the relative position of a context to the noise configuration is an important factor in spatial configuration learning, along with the relative position of the search items within the context (cf. Chua & Chun, 2003; Jiang & Wagner, 2004).

Olson and Chun (2002) showed that the magnitude of the contextual cuing effect when only half of the configuration was repeated was comparable with that when the whole configuration was repeated. Their finding suggests two possibilities about how the spatial context could be learned: One is that the local context is extracted from the noise information and is stored in memory independently of the noise information. The other is that the local context is encoded and stored with some type of noise information, such as the spatial relationship between the visual context and the noise. Regarding the latter hypothesis, the present study demonstrated that the contextual cuing effect disappeared when the relative position of visual context and noise was changed. This finding suggests that the visual context includes some kind of noise information, that at least the relative positions of the repeated and noise configurations was encoded and stored in memory.

Although there is no evidence that visual context in the contextual cuing paradigm is the same as that in natural scene perception, the results of the present study are consistent with the findings from studies of natural scene perception. For example, Biederman (1972) investigated the role of context, including object location and identity, in scene perception. He demonstrated that the accuracy of target identification was significantly lower when parts of the picture were jumbled (as in the shift condition in Experiment 1), even when participants had knowledge of the image to be presented. This finding suggests that the use of contextual information was affected when a scene had some kind of irregularity. The present finding that contextual cuing is restricted by the relative position of the learned context seems to be consistent with this result. The authors of other recent studies have proposed that the retention of visual short-term memory is also more affected by a change in the relative rather than the absolute position of the items in a display (Jiang, Olson, & Chun, 2000; Olson & Marshuetz, 2005). All of these findings imply that encoding relative position is a fundamental function of the human visual system.

REFERENCES

- BIEDERMAN, I. (1972). Perceiving real-world scenes. *Science*, **177**, 77-80.
- BIEDERMAN, I., MEZZANOTTE, R. J., & RABINOWITZ, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, **14**, 143-177.
- BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.

- CHUA, K.-P., & CHUN, M. M. (2003). Implicit scene learning is view-point dependent. *Perception & Psychophysics*, **65**, 72-80.
- CHUN, M. M. (2000). Contextual cuing of visual attention. *Trends in Cognitive Sciences*, **4**, 170-178.
- CHUN, M. M., & JIANG, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, **36**, 28-71.
- CHUN, M. M., & JIANG, Y. (2003). Implicit long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 224-234.
- CHUN, M. M., & NAKAYAMA, K. (2000). On the functional role of implicit visual memory for adaptive deployment of attention across scenes. *Visual Cognition*, **7**, 65-81.
- JIANG, Y., OLSON, I. R., & CHUN, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 683-702.
- JIANG, Y., & WAGNER, L. C. (2004). What is learned in spatial contextual cuing: Configuration or individual locations? *Perception & Psychophysics*, **66**, 454-463.
- OLSON, I. R., & CHUN, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, **9**, 273-302.
- OLSON, I. R., & MARSHUETZ, C. (2005). Remembering "what" brings along "where" in visual working memory. *Perception & Psychophysics*, **67**, 185-194.
- PELLI, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.

NOTES

1. Mean error rates were each subjected to a two-way ANOVA with epoch and condition as factors. However, no main effects or interactions were observed in either phase. It is therefore very unlikely that the RT effects were the result of differential speed-accuracy trade-offs.

2. The difference between the mean hit and false alarm rates was significant [$t(16) = 2.62, p < .05$]. To test the possibility that the contextual cuing effect in this experiment was due to explicit memory of the old configurations, we conducted an additional analysis. For the 16 configurations used in the *old* condition at test, we calculated the mean RTs for the hit and the miss configurations in each epoch and subjected them to a two-way ANOVA with the factors epoch (1-8) and recognition performance (hit or miss). The main effect of epoch was significant [$F(7,112) = 14.24, p < .001$], but neither the main effect of recognition performance nor the interaction was ($F_s < 1$). Thus, it is unlikely that the contextual cuing effect observed in this experiment depended on explicit memory.

3. Mean error rates were each subjected to a two-way ANOVA with epoch and condition as factors. However, no main effects or interactions were observed in either phase. It is therefore very unlikely that the RT effects were the result of differential speed-accuracy trade-offs.

(Manuscript received May 26, 2004;
revision accepted for publication February 3, 2005.)