Attention, Perception, & Psychophysics
2009, 71 (7), 1514-1524
doi:10.3758/APP.71.7.1514

Region segmentation and contextual cuing in visual search

Markus Conci
Ludwig-Maximilians-Universität München, Munich, Germany

and

Adrian von Mühlenen
University of Warwick, Coventry, England

Contextual information provides an important source for behavioral orienting. For instance, in the contextual-cuing paradigm, repetitions of the spatial layout of elements in a search display can guide attention to the target location. The present study explored how this contextual-cuing effect is influenced by the grouping of search elements. In Experiment 1, four nontarget items could be arranged collinearly to form an imaginary square. The presence of such a square eliminated the contextual-cuing effect, despite the fact that the square's location still had a predictive value for the target location. Three follow-up experiments demonstrated that other types of grouping abolished contextual cuing in a similar way and that the mere presence of a task-irrelevant singleton had only a diminishing effect (by half) on contextual cuing. These findings suggest that a segmented, salient region can interfere with contextual cuing, reducing its predictive impact on search.

A fundamental task of the human information processing system is to structure the complex input from the ambient visual array. Perceptual grouping provides one mechanism that integrates fragmentary parts into coherent units or objects while segregating it from other, neighboring objects and from the background (see Roelfsema, 2006, for a review). Although it is believed that image segmentation is primarily bottom-up driven, there are also studies suggesting that it can be influenced by previously acquired knowledge, for instance, by exploiting shape familiarity (e.g., Nelson & Palmer, 2007; Vecera & Farah, 1997) or the statistical covariation of objects (Baker, Olson, & Behrmann, 2004) and their spatial relations (Chun & Jiang, 1998). For example, in Chun and Jiang's study, learned contextual layouts facilitated target detection in subsequent trials. Thus, both grouping and contextual learning establish relational structures, either by means of bottom-up segmentation or by means of retrieving learned contextual associations. Here, we investigate how both processes, grouping and contextual learning, interact with each other.

Perceptual grouping has originally been shown to follow a set of basic principles in imposing structure on fragmentary information (Koffka, 1935). In many cases, grouped items allowed more efficient processing than did the corresponding processing of individual component parts. For instance, configural properties are processed more efficiently than individual features are (Pomerantz, Sager, & Stoever, 1977), and studies on visual search have shown that the individual parts may be integrated into a coherent representation prior to the engagement of attention (e.g., Moore & Egeth, 1997; Rensink & Enns, 1995). Also, search for a coherent object (composed from individual items/elements that can be grouped to form a global object representation) may be more efficient than is search for the same basic elements if these are presented such that no integrated object can be established (Conci, Müller, & Elliott, 2007b; Found & Müller, 1997; see also Friedman-Hill & Wolfe, 1995). In accordance with the basic Gestalt principles, low-level grouping mechanisms have been shown to structure visual input according to a variety of general laws, such as similarity (Duncan, 1984; Duncan & Humphreys, 1989), closure (Conci, Müller, & Elliott, 2007a; Donnelly, Humphreys, & Riddoch, 1991), and proximity (Han, Humphreys, & Chen, 1999). For instance, Conci et al. (2007a) demonstrated that visual search for a collinear line grouping is guided by the integrated representation of a closed shape. More specifically, search for a closed target configuration (composed from Ls arranged to form a square grouping) among open distractor configurations (arranged to form an open cross shape) was more efficient than was search for an open (cross-shaped) target configuration among closed square distractors, even though the local contrasts between elements making up the configurations were held constant. This suggests that closed objects primarily determine the way in which mechanisms of region segmentation identify salient units during search.

M. Conci, conci@psy.uni-muenchen.de

Apart from influences from low-level grouping operations, knowledge accumulated across previous trials may also have a significant impact on how a complex scene is decomposed. For example, Vecera and Farah (1997) showed that item familiarity determines image segmentation while influencing the exogenous orienting of attention (see also Nelson & Palmer, 2007). In addition, visual covariation can support the recognition of objects in complex displays (see Chun, 2000; Chun & Nakayama, 2000, for reviews). In the domain of visual search, Chun and Jiang (1998) showed that observers can learn the spatial relations between search items and use this in subsequent trials to guide attention more efficiently to the target. In their contextual-cuing experiment, search arrays consisted of 1 target T and 11 nontarget Ls (see Figure 1A for an example of a comparable display). Displays differed in that targets could appear in either an old or a new configuration: For old configurations, the target was always presented within the same arrangement of nontargets. These configurations were compared with new configurations that always presented novel nontarget arrangements on each trial. Consequently, the difference between old and new configurations indicates whether there is an influence of invariant spatial layouts (i.e., layouts that are repeated over and over) on the difficulty of target detection. The results, in fact, showed that the repetition of the spatial arrangement of a given old configuration led to a benefit in the mean reaction time (RT), as compared with when the

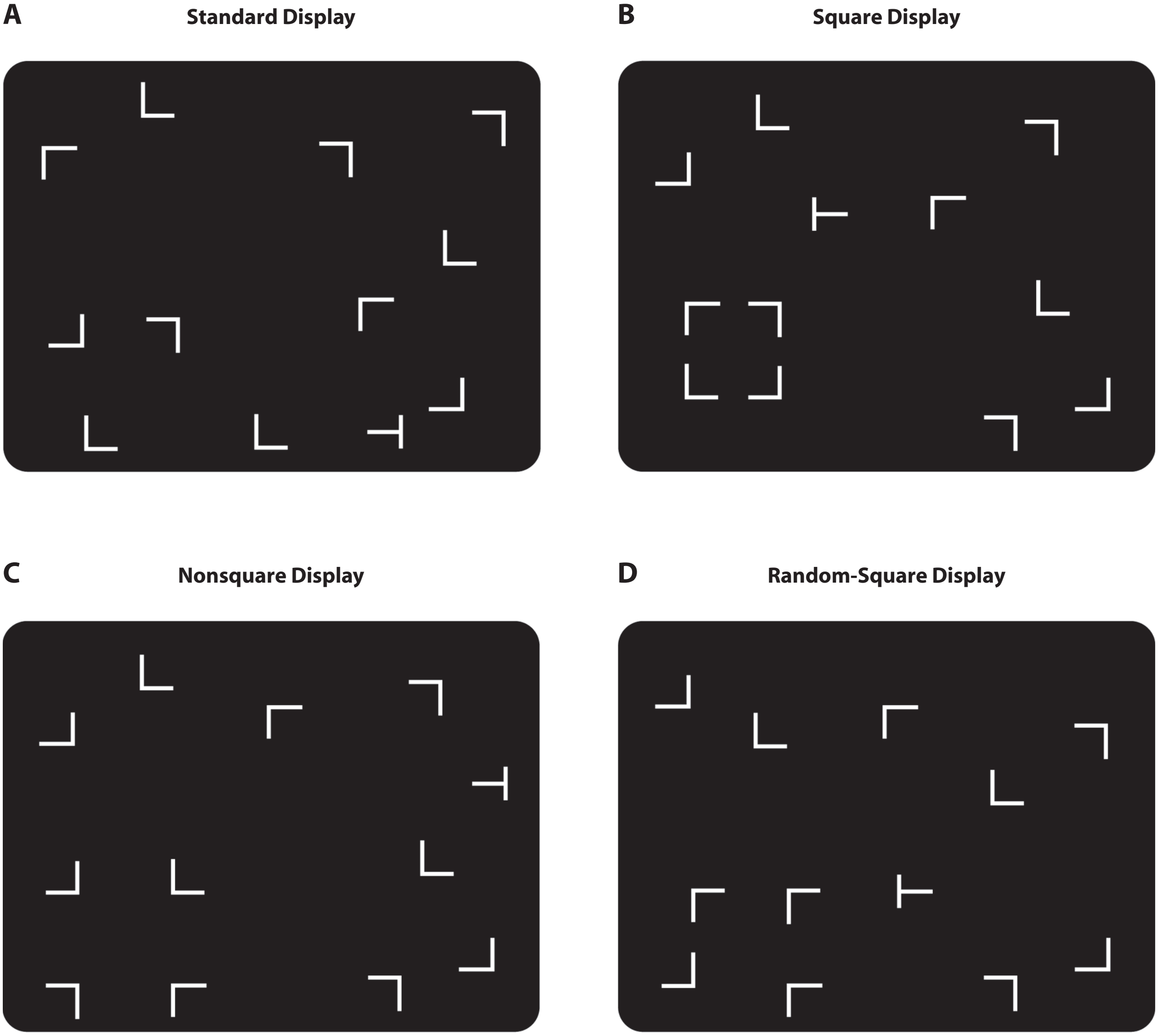

Figure 1. Examples of search displays: Each display contained one target (T) pointing to the left or to the right among 11 nontargets (Ls). The standard displays (A) served as the baseline in all experiments. For square displays (B), 4 of the 11 nontarget elements were arranged to form a collinear square grouping (Experiment 1). Similarly, for nonsquare displays (C), 4 nontargets were again presented in square-like form, but with the square elements pointing toward the center of the square configuration (Experiment 3). For random-square displays (D), 4 elements were presented in square-like form, but this time with random orthogonal orientations (Experiment 4).

spatial layout was new (the contextual-cuing effect), without observers being able to explicitly discern repeated displays from novel arrangements (but see Smyth & Shanks, 2008). These findings suggest that spatial associations are encoded, thus facilitating performance.

Both sets of findings, about the effects of grouping and contextual cuing, show that two distinct types of mechanisms could provide structure in order to decompose a visual scene efficiently: On one hand, bottom-up stimulus-driven perceptual grouping may provide a means to cluster salient regions within a given display layout (e.g., Conci et al., 2007a; Donnelly et al., 1991; Pomerantz et al., 1977). On the other hand, the relations between items of a scene are apparently encoded and remembered, guiding attention on the basis of the past experience (e.g., Chun & Jiang, 1998). It is, however, not clear if and in what way these two mechanisms interact with each other.

In the present study, we set out to explore the relation between perceptual grouping and memory-based guidance within the framework of contextual cuing. Four experiments investigated how perceptual grouping and attention can affect the learning of contextual information. Randomly generated displays similar to those used in previous studies served as a baseline measure in all experiments (see Figure 1A for an example). These standard displays were compared with randomly generated displays that either made up an "accidental" group of four close-by nontarget items forming a square arrangement (see Figure 1B for an example) or contained a single salient, red nontarget item. Our hypothesis was that the integration of elements into a salient group would significantly modulate the learning of contextual information: Given that contextual cuing has been shown to operate on a set of associative links established between the nontarget items and the target item (Brady & Chun, 2007), we expected that presenting a salient group of nontargets would provide a particularly strong link, allowing efficient target retrieval. Alternatively, if context-based memory representations were independent from the bottom-up display segmentation processes, then no effect of the grouped display layouts should be observable on the size of contextual cuing.

EXPERIMENT 1

Experiment 1 was conducted to investigate how the integration of four elements into a salient group affects the encoding of spatial context in visual search. We employed the contextual-cuing paradigm (Chun & Jiang, 1998) to compare how memory-based attentional guidance differs for displays that either contain a group of search items (square displays) or are without any specific groupings (standard displays; see Figure 1).

Method

Participants. Ten observers (4 male; mean age = 23.7 years) with normal or corrected-to-normal visual acuity participated in the experiment, receiving payment of €8/h. Participants were not aware of the purpose of the study.

Apparatus and Stimuli. The experiment was conducted on an IBM PC-compatible computer using MATLAB routines and Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli subtended 0.7º × 0.7º and were presented in gray (8.5 cd/m^2) against a black (0.02 cd/m^2) background on a 17-in. CRT monitor. A search display always consisted of 12 items, 1 target and 11 nontargets. The target was a T shape rotated 90º either clockwise or counterclockwise with random probability. Nontargets were L shapes rotated randomly in one of four orthogonal orientations. Search displays were generated by placing 1 T and 11 Ls randomly within the cells of an 8 × 6 matrix (cell size 2.5º). Within each cell, the positions of the stimuli were randomly jittered horizontally and vertically in steps of 0.1º within a range of ±0.6º. Two types of displays were generated. *Standard* displays had a randomized spatial layout without any specific item arrangements (see Figure 1A). In *square* displays, 4 of the 11 nontarget items were presented at adjacent (2 × 2) nonjittered cells, with L shapes rotated such that the four items formed a collinear square (subtending 2.9º × 2.9º). An example display is shown in Figure 1B. The location of the 2 × 2 cell square within the 8 × 6 cell matrix varied randomly among the 35 possible locations.

Trial sequence. Each trial started with the presentation of a central fixation cross for 500 msec. The fixation cross was followed by the search display, to which participants responded with a speeded response via mouse keys. 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. Displays remained on screen until a response was recorded. In case of an erroneous response, feedback was provided by an alerting sign ("−") presented for 1,000 msec at the center of the screen. The intertrial interval was 1,000 msec.

Design and Procedure. We used a three-factors within-subjects design with context, display type, and epoch as independent variables. Context had two levels: old and new. For the old-context condition, the arrangement of nontarget items was the same on every presentation. In the new-context condition, a new, random arrangement of nontarget items was generated on every presentation. In order to rule out location probability effects, the target appeared equally often at 24 possible locations throughout the experiment. The orientation of the target was determined randomly for each trial, whereas the orientations (and identities) of the nontarget items were preserved for the old-context condition. The second variable, display type, also had two levels: standard and square. In standard displays, all nontarget items were presented at randomly determined locations. In square displays, four nontarget items formed a square, whereas the other items were at randomly determined locations. Note that square displays in the old-context condition preserved the position of all nontarget items, including the square. Finally, the third variable, epoch, simply divided the experiment into six subsequent bins, allowing the assessment of possible learning effects over the course of the experiment.

At the beginning of the experiment, participants completed 1 block of 24 randomly generated practice trials to get familiarized with the task. All subsequent experimental blocks contained the same 12 old-context displays and 12 new-context displays in randomized order. In each block, old- and new-context displays were further subdivided into standard and square displays (i.e., 6 old- and 6 new-context standard displays and 6 old- and 6 new-context square displays). There were 30 blocks in the experiment, with 720 experimental trials in total.

Recognition test. After completing the search task, participants were asked to perform a recognition test. They were informed that certain display configurations had been repeated throughout the experiment and that they had to decide whether a given display had been shown previously. Of the 24 displays shown, 12 were old-context displays that were used in the experiment, and 12 were newly generated. The trial sequence was identical to that in the search task, except that no error feedback was given. Participants had to indicate whether the display was new or previously seen. Nonspeeded responses were recorded via left (new) and right (old) mouse keys.

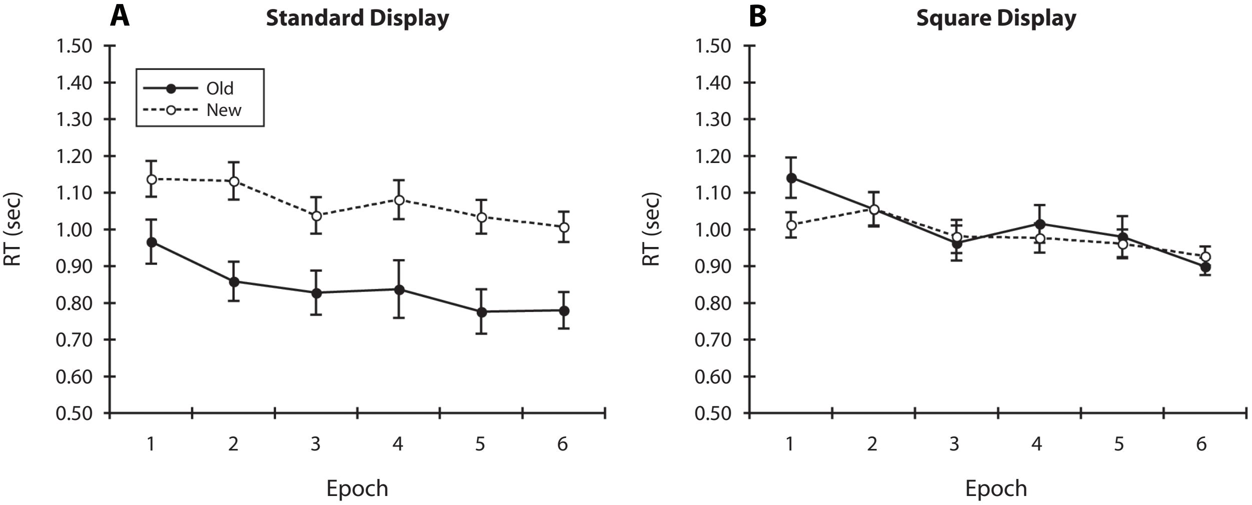

Figure 2. Mean reaction times (RTs; with standard error bars) as a function of epoch in the standard (A) and square (B) display conditions of Experiment 1. Filled and unfilled symbols correspond to old- and new-context conditions, respectively.

Results

Search task. Mean error rates were calculated for each participant and variable combination. The overall error rate was very low (1.0%), and a repeated measures ANOVA with the factors context (old, new), display type (standard, square), and epoch (1–6), revealed no significant effects (all *p*s $< .29$).

Individual mean RTs were computed for each variable combination, excluding erroneous responses and RTs greater than 3 sec. Fewer than 1% of all trials were excluded by this outlier criterion (which was also the case in all subsequent experiments). Figure 2 presents the mean correct RTs, averaged across participants, as a function of epoch for standard displays (left panel) and for square displays (right panel). Mean correct RTs were subjected to a three-way ANOVA with main terms for context (old, new), display type (standard, square), and epoch (1–6). The analysis revealed significant main effects of context [$F(1,9) = 8.63, p < .05$] and of epoch [$F(5,45) = 14.49, p < .001$]. Old-context trials were on average 104 msec faster than new-context trials, and search became faster with increasing epoch (RTs were 161 msec faster in Epoch 6 than in Epoch 1). The main effect of display type was only marginally significant [$F(1,9) = 3.90, p = .08$], indicating that RTs were overall a bit (41 msec) slower when a square was present. There was a significant interaction between context and epoch [$F(5,45) = 4.18, p < .01$], indicating that, overall, the contextual-cuing effect increased with proceeding epoch (from 83 msec in Epoch 1 to 167 msec in Epoch 6). Most interestingly, context interacted with display type [$F(1,9) = 40.20, p < .001$], indicating that the contextual-cuing effect occurred only with standard displays, but not with square displays (231 vs. −23 msec, respectively). This interpretation was confirmed by two split-up ANOVAs, where the main effect for context was significant only in the standard display ANOVA [$F(1,9) = 21.85, p < .001$], but not in the square display ANOVA [$F(1,9) = 0.63, p = .44$]. To further explore the onset of contextual learning with standard displays, mean RTs were calculated separately for each block.[1] As can be seen from Figure 3, the contextual-cuing effect emerged very early and was already evident in the third block. A series of *t* tests comparing old and new contexts confirmed this observation. The effect of contextual repetition was significant only from Block 3 onward (all *p*s $< .05$, except for Block 16). A similar pattern of rapid contextual learning was also observed in the subsequent experiments.

Recognition test. Overall, the mean accuracy in the recognition test was 54%. For standard displays, participants correctly identified old patterns on 64.9% of all trials (hit rate), but this differed only marginally from their false alarm rate of 50.1% [$t(9) = 2.17, p = .06$]. For square displays, the hit rate was 45.0% and the false alarm rate was 45.1%, indicating no significant differences [$t(9) = 0.01, p = .99$]. Thus, the marginal difference in standard displays indicates that participants were to a certain extent aware that some displays were repeated (see Smyth & Shanks, 2008, for a similar finding).

Discussion

The results for standard displays in Experiment 1 replicated previous findings on contextual cuing in visual search: Participants were significantly faster in detecting the target within old-context displays than within new-context displays, and this difference in performance increased in the course of the experiment. As in previous studies (e.g., Chun & Jiang, 1998; Lleras & von Mühlenen, 2004), the memorized contextual information could be used to guide spatial attention to the target location, leading to an RT benefit for old- versus new-context displays.

With square displays, where four nontarget items formed an imaginary square, no contextual-cuing effect was obtained. This is interesting, because the square itself might still have been predictive of the target location in

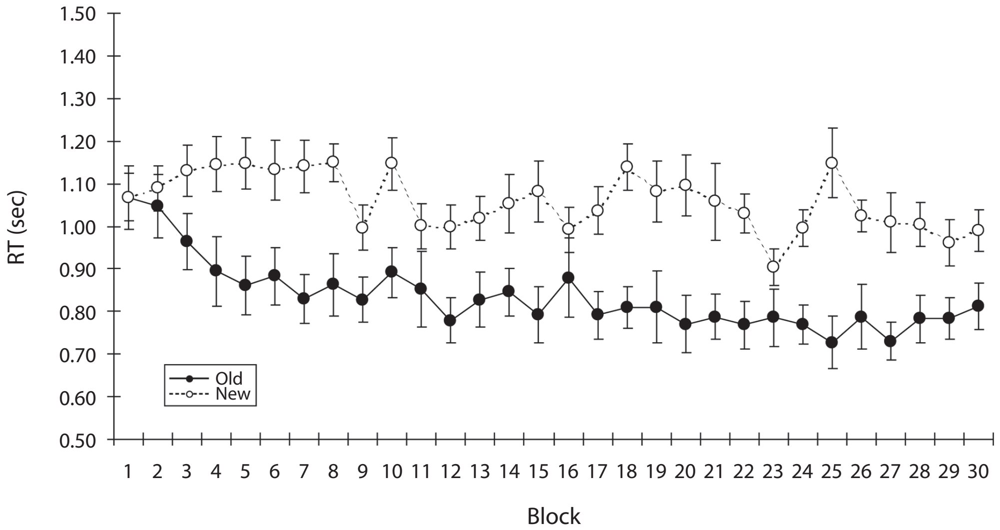

Figure 3. Mean reaction times (RTs; with standard error bars) as a function of block for the standard displays of Experiment 1. Filled and unfilled symbols correspond to old- and new-context conditions, respectively.

old-context displays, and, in fact, as outlined in the introductory section, our expectation was that the square would provide a particularly salient cue to retrieve the target. However, even though no contextual cuing was found with square displays, search performance overall benefited from the grouping of nontarget items. When only new-context trials are considered, mean RTs show that targets are found overall 86 msec faster in square displays than in standard displays [$t(9) = 2.89$, $p < .02$]. This indicates that, overall, the grouping operations had a beneficial effect on the search process, presumably by reducing the number of distracting items that needed to be inspected.

Despite increased overall search efficiency, presenting a square configuration in the display had the effect of abolishing contextual cuing, possibly because the salient, collinear grouping required attentional resources. For instance, a recent study by Yeshurun, Kimchi, Sha'shoua, and Carmel (2009) showed that a square grouping can attract attention, indicating that RT costs and RT benefits are to whether the target is located outside or within the square's boundaries. Likewise, for square displays, attentional resources were possibly not available to process and consequently to learn the context of the search displays.[2] In line with this view are previous studies reporting that the implicit learning of contextual information critically depends on the deployment of selective attention to predictive information (Jiang & Chun, 2001; Jiang & Leung, 2005). Thus, we hypothesized that, with square displays, contextual cuing did not occur, because the square acts as a salient cue that draws attentional resources away from the overall configuration of the displays.

EXPERIMENT 2

Would any salient stimulus disrupt contextual cuing? To investigate this possibility, in Experiment 2, we presented displays containing 1 search item in a salient red color. Such color singletons have been reported as being efficient in attracting attentional resources (e.g., Theeuwes, 1992). Thus, we tested whether it would be sufficient to present a single salient item in order to eliminate contextual cuing. Experiment 2 was identical to Experiment 1, except that the square displays were replaced by singleton displays, where 1 of the 11 nontarget items was red.

Method

Apparatus, stimuli, design, and procedure were identical to those in Experiment 1, except that square displays were replaced by singleton displays. Singleton displays were identical to standard displays except that one randomly selected nontarget item was colored in red (6.9 cd/m^2), whereas the remaining items were gray (8.5 cd/m^2). Red and gray items were subjectively matched in luminance by the experimenter. Note that the location of the red item in old displays was (like the square location in Experiment 1) preserved. All other details of the experiment were identical to those in Experiment 1. Ten observers (3 male; mean age = 29.3 years) with normal or corrected-to-normal visual acuity participated in the experiment, receiving course credits or €8/h.

Results

Search task. Erroneous responses were rare (1.0%), and the ANOVA with the factors context (old, new), display type (standard, singleton), and epoch (1–6) revealed no significant effects (all ps $> .28$).

Individual mean RTs were computed, excluding erroneous responses and outliers with RTs greater than 3 sec. Figure 4 presents the mean correct RTs averaged across participants as a function of epoch for standard displays (left panel) and singleton displays (right panel). Mean correct RTs were subjected to a three-way ANOVA with main terms for context, display type, and epoch. The analysis revealed significant main effects for context [$F(1,9) = 151.88$, $p < .001$] and display type [$F(1,9) = 18.69$, $p < .003$], whereas

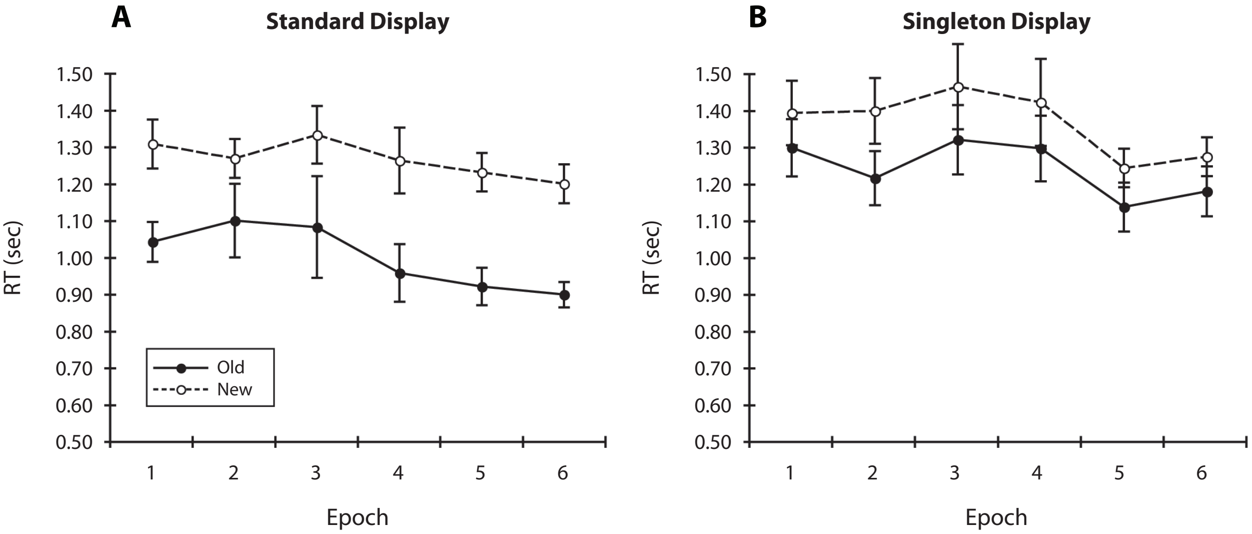

Figure 4. Mean reaction times (RTs; with standard error bars) as a function of epoch in the standard (A) and singleton (B) display conditions of Experiment 2. Filled and unfilled symbols correspond to old- and new-context conditions, respectively.

the main effect of epoch was only marginally significant [$F(5,45) = 2.34$, $p = .06$]. Old-context trials were on average 195 msec faster than new-context trials, and search in singleton displays was 169 msec slower than in standard displays. In addition, RTs became somewhat faster as the experiment progressed (RTs were 122 msec faster in Epoch 6 than in Epoch 1). Further, display type showed marginal two-way interactions with epoch [$F(5,45) = 2.28$, $p = .06$] and with context [$F(1,9) = 4.62$, $p = .06$]. The trend of a display type $\times$ epoch interaction depicted, for standard displays, a decrease in the mean RTs with epoch, whereas for singleton displays, the response latencies were more variable across the experiment (see Figure 4). In addition, the marginal display type $\times$ context interaction indicated that the contextual-cuing effect was smaller for singleton displays (124 msec) than for standard displays (267 msec). However, two split-up ANOVAs showed that the main effect for context was significant in both the standard display ANOVA [$F(1,9) = 39.01$, $p < .001$] and the singleton display ANOVA [$F(1,9) = 17.77$, $p = .003$], indicating that there was a robust contextual-cuing effect in both standard and singleton display conditions. Finally, the three-way interaction was significant [$F(5,45) = 2.54$, $p < .05$]. Further explorations of this interaction revealed that it was mainly due to Epoch 2, where standard displays showed (in comparison with the other epochs) a relatively weak contextual-cuing effect (168 vs. 286 msec, respectively) and singleton displays showed a relatively strong contextual-cuing effect (183 vs. 113 msec).

Recognition test. The overall mean accuracy in the recognition test was 57%. For standard displays, participants correctly identified old patterns on 53.3% of all trials (hit rate), but this did not differ from their false alarm rate of 44.8% [$t(9) = 0.76$, $p = .46$]. For singleton displays, the hit rate was 61.6% and the false alarm rate was 45.0%, again indicating no significant differences [$t(9) = 1.46$, $p = .17$]. Thus, the lack of a significant recognition-test difference in Experiment 2 suggests that observers were not explicitly aware of the display repetitions (despite a clear numerical trend in comparisons of hit and false alarm rates).

Discussion

Experiment 2 revealed for standard displays the same robust contextual-cuing effect as in Experiment 1 (267 vs. 231 msec). However, the singleton displays did not have the same effect on contextual cuing as the square groupings did: Although context effects were reduced to a certain extent, they still showed a facilitatory influence of 124 msec on the RTs. In comparison, presenting a square in Experiment 1 had the effect that contextual cuing essentially disappeared, revealing only a small difference of −23 msec between old and new displays.

Although the contextual-cuing effect was less affected by the singleton than by the square, a direct comparison of the new-context conditions indicated that the presence of a salient red item led to reliable RT costs of 94 msec relative to the standard display condition [$t(9) = 2.85$, $p < .02$]. This slowing of RTs is in line with Theeuwes's (1992) finding that a salient color singleton delays the response latency, which had been taken as an indicator for attentional capture. However, the occurrence of attentional capture does not seem sufficient to entirely eliminate the contextual-cuing effect, since repetitions of the singleton displays still led to substantial RT benefits. Therefore, the attentional distraction away from the context toward a singleton cannot fully account for the complete absence of contextual cuing with square configurations. One possibility could be that the color singleton was simply less effective than the square was at capturing attention (for instance, the effects go in the predicted direction, showing a reduction of contextual cuing). However, we believe that this is unlikely, in that the red singleton led to significant (94-msec) capture. Instead, as an alternative, the lack of contextual cuing in square displays in Experiment 1 suggests that other factors, such as grouping the Ls into a coherent square, abolish the RT facilitation provided by the contextual repetitions.

EXPERIMENT 3

The results of Experiment 2 suggest that the eliminating effect of the square on contextual cuing in Experiment 1 was probably to some degree due to attentional capture by the square. Experiment 3 further investigated whether the other part of the eliminating effect on contextual cuing in Experiment 1 was due to interference from grouping processes involved in binding the four Ls forming the square. The displays were the same as in Experiment 1, except that the four L shapes forming the figure were each turned by 180º to form a nonsquare group. That is, the corner junctions of the four Ls now pointed toward the center of the group, forming the open shape of an imaginary cross (see Figure 1C for an example display). Consequently, although this nonsquare group lacked closure, it still exhibited a symmetric organization with aligned corner junctions, which was expected to produce a strong grouping. Due to its openness, the elements of the nonsquare were more embedded in the overall search display, making the figure less prominent. It was therefore expected that the nonsquare would not, or would to a much lesser extent, attract attention than the square would. Evidence for this idea also comes from Conci et al. (2007a), who showed that visual search for an open shape (like the one in Figure 1A) among closed shapes (like the ones in Figure 2C) is less efficient than vice versa (a search for a closed shape among open shapes), suggesting that open shapes are less attention grabbing than closed shapes are. Consequently, if the lack of contextual cuing in the square displays in Experiment 1 resulted, in part, from perceptual grouping (and not only from attentional capture), then the cross-shaped (nonsquare) grouping should also interfere to some extent with contextual cuing (despite the reduced saliency of the figure).

Method

Apparatus, stimuli, design, and procedure were identical to those in Experiment 1, except that square displays were replaced by nonsquare displays. Nonsquare displays consisted of 12 items, with one cluster of four nontargets presented next to each other and with the L shapes rotated, such that corner junctions pointed toward the center of the configuration. The resulting grouping integrated individual elements into a symmetric cross shape (centered around an invisible square with side lengths of 2.9º). All other details of the experiment were identical to those in Experiment 1. Ten observers (4 male; mean age = 31.8 years) with normal or corrected-to-normal visual acuity participated in the experiment, receiving €8/h.

Results

Search task. Erroneous responses were rare (1.0%), and the ANOVA with the factors context (old, new), display type (standard, nonsquare), and epoch (1–6) revealed only a marginally significant three-way interaction [$F(5,45) = 2.16$, $p = .08$; all other ps $> .34$].

Individual mean RTs were computed, excluding erroneous responses and outliers with RTs greater than 3 sec. Figure 5 presents the mean correct RTs averaged across participants as a function of epoch for standard displays (left panel) and nonsquare displays (right panel). Mean correct RTs were subjected to a three-way ANOVA with main terms for context, display type, and epoch. The analysis revealed significant main effects for context [$F(1,9) = 111.74$, $p < .001$], display type [$F(1,9) = 69.48$, $p < .001$], and epoch [$F(5,45) = 10.61$, $p < .001$]. Old-context trials were on average 159 msec faster than new-context trials, RTs became faster as the experiment progressed (RTs were 121 msec faster in Epoch 6 than in Epoch 1), and search in nonsquare displays was 157 msec slower than in standard displays. Further, context interacted with display type [$F(1,9) = 49.47$, $p < .001$]. As was the case in Experiment 1, the contextual-cuing effect occurred only with standard displays and not with nonsquare displays (345 vs. −28 msec, respectively). This interpretation was confirmed by two split-up ANOVAs,

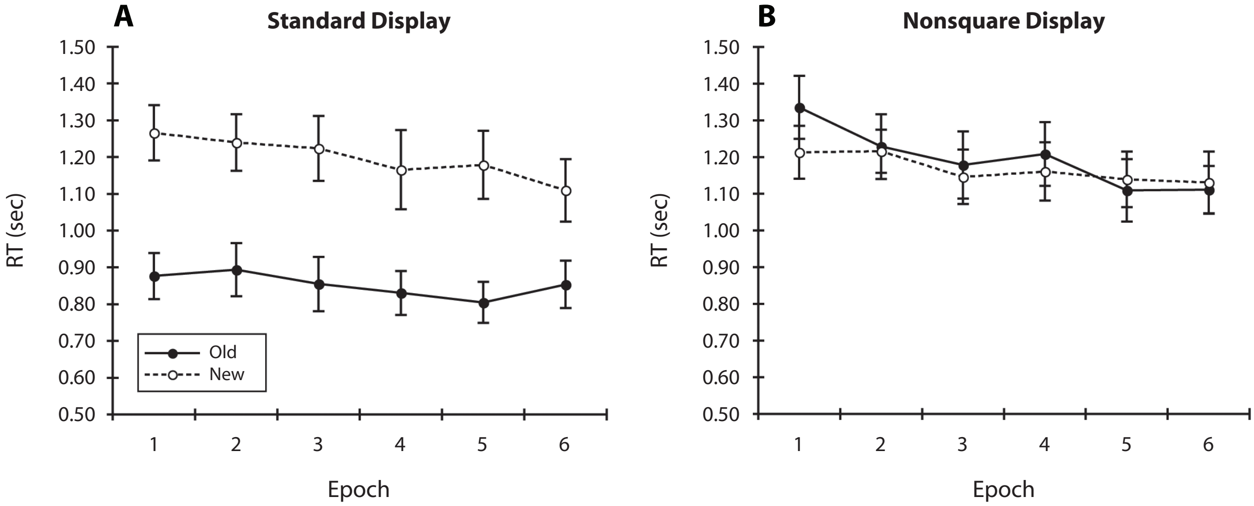

Figure 5. Mean reaction times (RTs; with standard error bars) as a function of epoch in the standard (A) and nonsquare (B) display conditions of Experiment 3. Filled and unfilled symbols correspond to old- and new-context conditions, respectively.

where the main effect for context was significant only in the standard [$F(1,9) = 119.14$, $p < .001$] and not in the nonsquare [$F(1,9) = 0.77$, $p = .41$] display ANOVA. Finally, the overall ANOVA also revealed a significant three-way interaction [$F(5,45) = 4.06$, $p < .05$]. The two split-up ANOVAs above indicate that this was due to a significant context × epoch interaction with standard displays [$F(4,45) = 5.36$, $p < .001$], which was not significant with nonsquare displays [$F(5,45) = 0.92$, $p = .47$]. For standard displays, the contextual-cuing effect became somewhat smaller as the experiment progressed from 389 msec in Epoch 1 to 256 msec in Epoch 6, whereas no significant modulation was obtained for the nonsquare displays. Note, however, that despite this decrease, contextual cuing still produced a strong and reliable RT benefit in all epochs.

Recognition test. The overall mean accuracy in the recognition test was 54%. For standard displays, participants correctly identified old patterns on 58.5% of all trials (hit rate), but this did not differ from their false alarm rate of 51.7% [$t(9) = 0.82$, $p = .43$]. For nonsquare displays, the hit rate was 48.3% and the false alarm rate was 38.3%, again indicating no significant differences [$t(9) = 1.41$, $p = .19$]. Thus, in Experiment 3, observers were not explicitly aware of the display repetitions (even though there was a certain numerical advantage for the hit rates relative to the false alarm rates).

Discussion

Experiment 3 mirrored the findings of Experiment 1, showing that a robust contextual-cuing effect completely disappears when four nontarget items form a nonsquare figure. This shows that a relatively nonsalient grouping (which is not defined by closure) disrupts contextual cuing as much as a salient group does. Thus, a simple aligned and symmetric arrangement of search items is sufficient to interfere with contextual cuing.

In contrast to Experiment 1, search performance in Experiment 3 was not affected by the presence of nonsquare objects in the new-context condition [$t(9) = 0.73$, $p = .48$]. This is in line with Conci et al.'s (2007a) finding that search for an open nonsquare is more difficult than one for a closed square (see also Donnelly et al., 1991). It further supports our assumption that the four elements of the nonsquare are seen as part of the search display. Mirroring these findings, between-experiment comparisons revealed that, in Experiment 1, the salient grouping into a square facilitated overall search performance, whereas, in Experiment 3, the reduction of saliency of the grouped nonsquare did not lead to a comparable benefit. This was also confirmed by an RT comparison for displays containing groups in the new-context condition of Experiments 1 and 3, which revealed that search was 182 msec slower for nonsquare displays than for square displays [$t(18) = 2.33$, $p < .04$]. However, the absence of a contextual-cuing effect in the presence of a figure in both experiments, is similar in both Experiments 1 and 3 [−23 and −27 msec, respectively; $t(18) = 0.13$, $p = .89$]. Thus, although the saliency of element groupings appears to influence the overall search performance, no comparable effect is depicted for the context-based learning mechanisms. Together with Experiments 1 and 2, the present experiment suggests that attention cannot be the sole factor influencing contextual cuing. Instead, our results strengthen the view that relatively subtle regularities in the spatial organization of the display layout that involve grouping processes are sufficient to drastically reduce the influence of contextual cuing.

EXPERIMENT 4

Experiments 1–3 showed that the contextual-cuing effect can be severely disrupted by regularities in the spatial organization of the display layout: When four nontarget items were grouped to form a symmetric cluster of elements, contextual cuing did not occur. Experiment 4 was performed to investigate what it takes to reinstate the contextual-cuing effect. Is grouping by collinearity, closure, or symmetry essential to disrupting the contextual-cuing effect, or is the simple clustering of four nontarget items (i.e., the proximal arrangement of four nontarget items) sufficient to disrupt contextual cuing? In Experiment 4, four nontarget items were presented as a random-square cluster, excluding symmetric arrangements (see Figure 1D for an example display).

Method

Experiment 4 was identical to Experiment 1, except that the square displays were replaced by random-square displays. Random-square displays consisted of 12 items with one cluster of four nontargets presented next to each other. The L shapes of this cluster were rotated at random such that each item could be presented in one of its four orthogonal orientations. Note that, although square and nonsquare groupings were not allowed as random squares, other regularities, such as the repetition of item orientations or the alignment of several elements, were allowed. All other details were identical to those in Experiment 1. Ten observers (5 male; mean age = 24.3 years) with normal or corrected-to-normal visual acuity participated in the experiment, receiving payment of €8/h.

Results

Search task. Erroneous responses were rare (1.6%), and an ANOVA with the factors context (old, new), display type (standard, random square), and epoch (1–6) revealed only a marginally significant effect of display type [$F(1,9) = 3.98$, $p = .08$; all other ps $> .39$].

Individual mean RTs were computed, excluding erroneous responses and RTs greater than 3 sec. Figure 6 presents the mean correct RTs averaged across participants as a function of epoch for standard displays (left panel) and random-square displays (right panel). A three-way ANOVA with main terms for context, display type, and epoch revealed a significant main effect of context [$F(1,9) = 37.44$, $p < .001$] and a marginally significant effect of epoch [$F(5,45) = 2.27$, $p = .06$]: Old-context trials were on average 144 msec faster than new-context trials, and RTs became 24 msec faster in Epoch 6 than in Epoch 1. In addition, there was a significant interaction between context and epoch [$F(5,45) = 2.42$, $p < .05$], indicating that contextual cuing increased a bit with time (from 111 msec in Epoch 1 to 148 msec in Epoch 6). Most

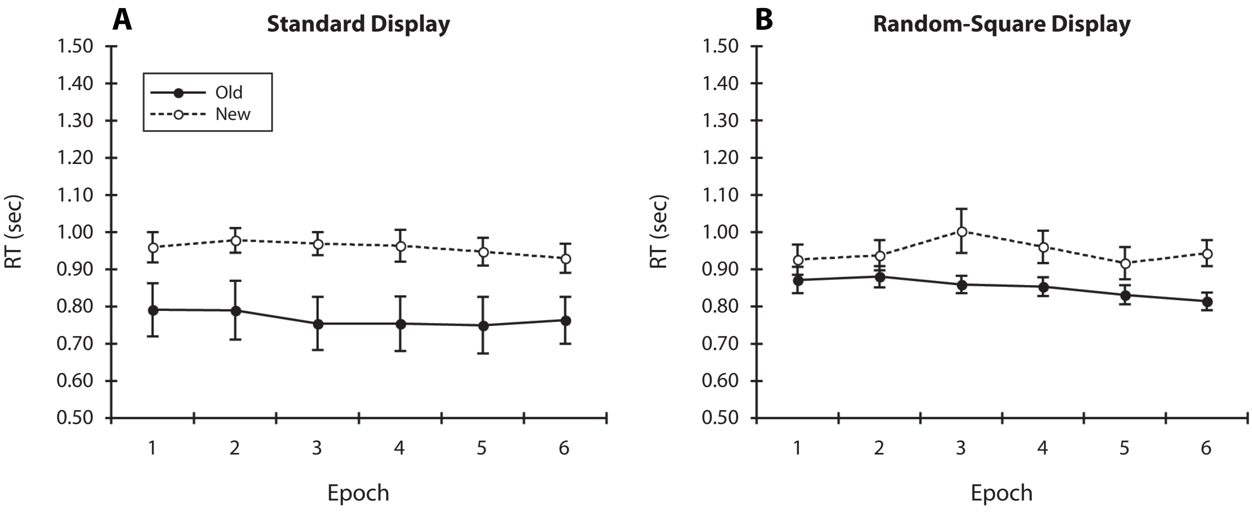

Figure 6. Mean reaction times (RTs; with standard error bars) as a function of epoch in the standard (A) and random-square (B) display conditions of Experiment 4. Filled and unfilled symbols correspond to old- and new-context conditions, respectively.

interestingly, there was no interaction between context and display type [$F(1,9) = 2.05$, $p = .19$], and a split-up ANOVA on random-square displays now showed a significant main effect for context [$F(1,9) = 9.84$, $p < .05$], indicating that contextual cuing was operating in both display type conditions.

Recognition test. The overall mean accuracy in the recognition test was 50%. For standard displays, participants correctly identified old patterns on 58.3% of all trials (hit rate), but this differed only marginally from their false alarm rate of 41.8% [$t(9) = 1.92$, $p = .09$]. For random-square displays, the hit rate was 48.3% and the false alarm rate was 65.0%, indicating again a marginally significant difference [$t(9) = 2.04$, $p = .07$]. Thus, participants were to a certain extent aware that some displays were repeated. However, importantly, the differences did not reach statistical significance.

Discussion

Unlike the square displays in Experiment 1 and the nonsquare displays in Experiment 3, the random-square displays did not abolish the contextual-cuing effect. This means that the proximal but random clustering of elements is not sufficient to interfere with contextual cuing, whereas item groupings that depict closure and collinearity (Experiment 1) or symmetry (Experiments 1 and 3) do interfere with contextual cuing. Without some figural regularity, a cluster with the proximal arrangement of four nontarget items does not interfere with cuing by contextual layouts.

Even though the display type $\times$ context interaction did not reach significance, Figure 6 suggests that contextual cuing was more pronounced with standard displays than with random-square displays (191 vs. 96 msec, respectively). One possibility to account for this reduction could be that, in a few displays, some of the random-square elements might still be grouped (e.g., by means of collinearity) into a partial object.

GENERAL DISCUSSION

The present set of experiments investigated how figural grouping and attention interact with contextual cuing. Our study showed that relatively subtle regularities, established by means of perceptually organized clusters among elements of a given display layout, had a profound influence on the contextual-cuing effect: In all experiments, a robust contextual-cuing effect was obtained with the standard displays. However, if four (task-irrelevant) nontarget items were grouped to form a collinear square, no contextual cuing occurred. By contrast, contextual cuing still showed a (somewhat reduced) facilitatory effect when a salient, red nontarget singleton was presented. Conversely, presenting a symmetric cross shape disrupted contextual cuing (to the same extent as did presenting a square) even though the cross-shaped configuration did not produce a closed shape and the overall search performance did not significantly benefit from nontarget grouping operations. Finally, if the four nontarget items were simply presented next to each other, contextual cuing occurred, comparable in magnitude to that for the standard displays.

Taken together, the experiments reported here show that memory-based contextual associations can be strongly influenced by accidental regularities in the displays. However, these regularities do not speed the response (as initially expected) but seem to disrupt contextual cuing. Presenting a configuration of nontarget items that group in terms of closure or symmetry leads to a drastic reduction of the contextual-cuing effect. Bottom-up perceptual groups therefore appear to govern search at the expense of memory-based contextual associations.

One way to interpret this finding would be that any salient group of elements (or any single element) attracts attention, diverting it away from the overall context of the search display. As a consequence of this lack of attention devoted to the display layouts, the contexts are simply not learned. In fact, attentional capture of salient square groupings has been documented in a recent study that showed that attention was automatically deployed to the grouping, interfering with target detection at other locations (Yeshurun et al., 2009). However, several observations speak against this interpretation: In Experiment 2, presenting a salient, red singleton clearly attracted attentional resources (as in other studies; e.g., Theeuwes, 1992). Nevertheless, the unique red item did not reveal a strong negative influence on contextual cuing (that would be comparable to the reduction observed for the square). Instead, display repetitions were reduced but still showed a substantial benefit on detection performance despite attentional capture. Conversely, in Experiment 3, the cross-shaped configurations did not affect the efficiency of search (see also Conci et al., 2007a) but, nevertheless, had a profound detrimental influence on contextual cuing. Consequently, this pattern of effects shows that attentional capture by an irrelevant stimulus can influence contextual cuing to some extent. However, our results also argue against attention's taking the sole role of modulating learning in contextual cuing (see also Jiang & Chun, 2001; Jiang & Leung, 2005; Rausei, Makovski, & Jiang, 2007).

A variant to explaining the lack of contextual-cuing effects when attention is distracted could be that learning itself was not affected, but that the recall of the learned context was (as proposed by Jiang & Leung, 2005). Thus, it could be that learning of *all* old display layouts was equally effective. However, difficulties arose when the learned contextual information was recalled, because attention was automatically drawn to the salient objects when the display was presented (and as a consequence, the contextual associations were lost). Nevertheless, again this explanation is unlikely given that, in Experiment 2, contextual cuing facilitated performance even though a salient item was present (that effectively captured attention).

Finally, as a potential alternative to an explanation in terms of attentional capture, the lack of support by contextual cuing in square and nonsquare displays could be explained with interference between perceptual grouping and contextual cuing. For instance, if one assumes that contextual learning is established primarily through a limited set of (three to four) target–nontarget associations (Brady & Chun, 2007; Song & Jiang, 2005), then learning in a random-display layout would provide a richer source of informative cues than would learning in displays that contain a subset of regular items that potentially lead to a bias in learning the context between the salient group and the target. Thus, in a display that contains a square grouping, context-based learning will preferably operate between the target and the grouped configuration. As a consequence, learning such a relation is less informative than the relation between the target and three or four spatially distributed nontargets, because finding a single object in a particular location is less of a suspicious coincidence than finding three or four objects in a particular location. Therefore, a certain degree of figural regularity will be sufficient to produce a sustained disruption of contextual cuing because the groupings provide a reduction in the variance of contextual information.

Modulations of contextual cuing with systematic changes in the spatial variance of nontarget items have also been reported in other studies. For instance, Olson and Chun (2002) showed that items that are spatially proximal to the target have a stronger impact for contextual cuing than the spatially segregated items do. In agreement with this finding, Brady and Chun (2007) reported that local context associations are particularly sensitive to contextual cuing. Similarly, for displays that are segregated in depth, contextual cuing is restricted within the local plane that contains the target (Kawahara, 2003). Although the individual associations between target and nearby nontargets appear to play a primary role in context-based learning, the global configuration has been shown to also contribute to contextual cuing (Jiang & Wagner, 2004). Display-irrelevant global background attributes have, however, only little influence on contextual cuing (Kunar, Flusberg, & Wolfe, 2006). Taken together, the overall evidence suggests that contextual cuing is primarily modulated by learned associations between the target and the surrounding items in a given display. However, if nontarget items form a strong group, then the association between the context and the target vanishes.

In sum, the present study suggests that attentional capture modulates the strength of contextual cuing, whereas segmented regions prevent the efficient learning of the display context. Grouping has, in many cases, been shown to play a major role in segmenting individual items into salient regions that serve as basic units for attentional processing (e.g., Conci et al., 2007a; Moore & Egeth, 1997; Pomerantz et al., 1977; Rensink & Enns, 1995; Wang, Kristjánsson, & Nakayama, 2005). In light of this bias toward salient groups, context-based learning is severely disrupted by processes of region segmentation.

AUTHOR NOTE

This work was supported by a German Research Foundation (DFG) Project Grant (FOR 481). We thank Barbara Neumayer and Philipp Schiebler for their help with data collection and Jeremy Wolfe, Liqiang Huang, and three anonymous reviewers for valuable comments on an earlier draft of the manuscript. Address correspondence to M. Conci, Allgemeine und Experimentelle Psychologie, Department Psychologie, Ludwig-Maximilians-Universität, Leopoldstrasse 13, D-80802 München, Germany (e-mail: conci@psy.uni-muenchen.de).

REFERENCES

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, **15**, 460-466.

Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception & Performance*, **33**, 798-815.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.

Chun, M. M. (2000). Contextual cueing 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., & Nakayama, K. (2000). On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Visual Cognition*, **7**, 65-81.

Conci, M., Müller, H. J., & Elliott, M. A. (2007a). Closure of salient regions determines search for a collinear target configuration. *Perception & Psychophysics*, **69**, 32-47.

Conci, M., Müller, H. J., & Elliott, M. A. (2007b). The contrasting impact of global and local object attributes on Kanizsa figure detection. *Perception & Psychophysics*, **69**, 1278-1294.

Donnelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 561-570.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, **113**, 501-517.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.

Found, A., & Müller, H. J. (1997). Local and global orientation in visual search. *Perception & Psychophysics*, **59**, 941-963.

Friedman-Hill, S., & Wolfe, J. M. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 531-551.

Han, S., Humphreys, G. W., & Chen, L. (1999). Parallel and competitive processes in hierarchical analysis: Perceptual grouping and encoding of closure. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1411-1432.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology*, **54A**, 1105-1124.

Jiang, Y., & Leung, A. W. (2005). Implicit memory of ignored visual context. *Psychonomic Bulletin & Review*, **12**, 100-106.

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cueing—Configuration or individual locations? *Perception & Psychophysics*, **66**, 454-463.

Kawahara, J. (2003). Contextual cueing in 3D layouts defined by binocular disparity. *Visual Cognition*, **10**, 837-852.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt.

Kunar, M. A., Flusberg, S. J., & Wolfe, J. M. (2006). Contextual cuing by global features. *Perception & Psychophysics*, **68**, 1204-1216.

Lleras, A., & von Mühlenen, A. (2004). Spatial context and top-down strategies in visual search. *Spatial Vision*, **17**, 465-482.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence for grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 339-352.

Nelson, R. A., & Palmer, S. E. (2007). Familiar shapes attract attention in figure–ground displays. *Perception & Psychophysics*, **69**, 382-392.

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, **9**, 273-302.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.

Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 422-435.

Rausei, V., Makovski, T., & Jiang, Y. V. (2007). Attention dependency in implicit learning of repeated search context. *Quarterly Journal of Experimental Psychology*, **60**, 1321-1328.

Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, **102**, 101-130.

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience*, **29**, 203-227.

Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory & Cognition*, **36**, 403-415.

Song, J.-H., & Jiang, Y. (2005). Connecting the past with the present: How do humans match an incoming visual display with visual memory? *Journal of Vision*, **5**, 322-330.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, **51**, 599-606.

Vecera, S. P., & Farah, M. J. (1997). Is visual image segmentation a bottom-up or an interactive process? *Perception & Psychophysics*, **59**, 1280-1296.

Wang, D. L., Kristjánsson, A., & Nakayama, K. (2005). Efficient visual search without top-down or bottom-up guidance. *Perception & Psychophysics*, **67**, 239-253.

Yeshurun, Y., Kimchi, R., Sha'shoua, G., & Carmel, T. (2009). Perceptual objects capture attention. *Vision Research*, **49**, 1329-1335.

NOTES

1. Note that these mean values are based on only six observations per participant.

2. Note that attentional capture should, in principle, decrease search efficiency, resulting in slower RTs, whereas the effect of grouping demonstrated here expedited RTs. One could, however, assume that the negative effect of capture was weaker than the positive effect of grouping.

(Manuscript received September 30, 2008;
revision accepted for publication May 23, 2009.)