

# The Nesting of Search Contexts Within Natural Scenes: Evidence From Contextual Cuing

Daniel I. Brooks, Ian P. Rasmussen, and Andrew Hollingworth  
University of Iowa

In a contextual cuing paradigm, we examined how memory for the spatial structure of a natural scene guides visual search. Participants searched through arrays of objects that were embedded within depictions of real-world scenes. If a repeated search array was associated with a single scene during study, then array repetition produced significant contextual cuing. However, expression of that learning was dependent on instantiating the original scene in which the learning occurred: Contextual cuing was disrupted when the repeated array was transferred to a different scene. Such scene-specific learning was not absolute, however. Under conditions of high scene variability, repeated search arrays were learned independently of the scene background. These data suggest that when a consistent environmental structure is available, spatial representations supporting visual search are organized hierarchically, with memory for functional subregions of an environment nested within a representation of the larger scene.

*Keywords:* scene perception, visual search, visual memory, spatial memory, contextual cuing

Intelligent behavior requires efficiently selecting goal-relevant objects in the environment. For example, when making coffee, attention must be directed sequentially to each object (coffee pot, cup, milk, spoon) that is needed as the activity unfolds (Land & Hayhoe, 2001). In the course of completing a real-world task, if the relevant object is not currently attended, a visual search operation is necessary to orient attention and the eyes to that object. If the coffee needs milk, attention must be directed efficiently to the milk (and not to the other objects in the kitchen), such that the milk carton can be grasped and the milk poured. This type of visual search operation occurs almost constantly during real-world tasks, and the efficiency by which objects are located directly influences the efficiency by which we are able to accomplish those tasks. Finding objects in complex, real-world environments is therefore a fundamental problem the visual system must solve, and understanding the means by which objects are located in scenes is central to understanding how vision supports goal-directed behavior.

The efficiency of visual search in scenes depends on multiple factors. Clearly, many of the factors traditionally studied in the literature on visual search are likely to influence search in natural environments, such as set size (Neider & Zelinsky, 2008), target-distractor similarity (Pomplun, 2006), and various other factors falling under the broad umbrella of target “salience” (Itti & Koch, 2000; Parkhurst, Law, & Niebur, 2002). However, the influence of

such relatively low-level visual features during real-world search is likely to be quite small compared with the influence of knowledge and memory (Becker & Rasmussen, 2008; Chen & Zelinsky, 2006; Eckstein, Drescher, & Shimozaki, 2006; Henderson, Brockmole, Castelano, & Mack, 2007; Henderson, Weeks, & Hollingworth, 1999; Hollingworth, 2009; Neider & Zelinsky, 2006; Torralba, Oliva, Castelano, & Henderson, 2006; Yarus, 1967). We often construct and organize environments precisely for the purpose of allowing memory and knowledge to govern visual search. The spoons tend to be kept in a particular drawer in the kitchen so that one need only inspect a single drawer to find one. After using a stapler, one tends to place it back in the same location on the desk so that memory for its location can guide attention back to the stapler the next time some stapling needs to be done. Even in an unfamiliar kitchen or office, one can use cultural knowledge of the typical locations of objects to restrict search to plausible regions of the scene, looking first in the kitchen drawers (rather than in the refrigerator) for a spoon or looking first on top of the desk (rather than underneath it) for a stapler.

Consistent with these observations, participants in laboratory experiments use knowledge about the category of a scene to limit search to regions of the scene where a particular type of target object is likely to be found (Eckstein et al., 2006; Henderson et al., 1999; Neider & Zelinsky, 2006; Torralba et al., 2006), such as searching for pedestrians in the lower regions of a street scene (Torralba et al., 2006). Memory for the structure of a individual scene exemplar also allows participants to direct attention and the eyes to regions of that scene that are likely to contain a particular search target (Castelano & Henderson, 2007; Hollingworth, 2009). Finally, memory for the specific location of the target object within a scene, acquired either from a preview of a scene (Hollingworth, 2009) or from repeated search through a scene (Becker & Rasmussen, 2008; Brockmole, Castelano, & Henderson, 2006; Brockmole & Henderson, 2006b), can guide attention efficiently to the target location.

---

This article was published Online First August 23, 2010.

Daniel I. Brooks, Ian P. Rasmussen, and Andrew Hollingworth, Department of Psychology, University of Iowa.

The study was supported by National Institutes of Health Grant R01EY017356. Aspects of this research were presented at the Eight Annual Meeting of the Vision Sciences Society, Sarasota, FL, May 2008.

Correspondence concerning this article should be addressed to Daniel I. Brooks or Andrew Hollingworth, University of Iowa, Department of Psychology, 11 Seashore Hall E, Iowa City, IA 52242-1407. E-mail: daniel-brooks@uiowa.edu or andrew-hollingworth@uiowa.edu

Although search through natural environments is typically governed by memory and knowledge, relatively little is understood about the memory representations of natural scenes guiding visual search. The nature of such spatial memory representations has been examined most directly in the literature on *contextual cuing* (Chun & Jiang, 1998; for a review, see Chun & Turk-Browne, 2008). In contextual cuing experiments, participants perform a series of searches through arrays of relatively simple stimuli (e.g., rotated *Ts* and *Ls*). Unbeknownst to the participant, the spatial configuration of some of the search arrays is repeated throughout the experiment. Search is facilitated as participants learn the spatial relationship between repeated search contexts and target locations. Similar results are obtained for repeated search through depictions of real-world scenes (Brockmole et al., 2006; Brockmole & Henderson, 2006a, 2006b). Contextual cuing has been interpreted as engaging the same statistical learning mechanisms responsible for learning contextual regularities in real-world environments (Chun & Turk-Browne, 2008) and to depend on medial temporal brain regions implicated in spatial–environmental learning (Chun & Phelps, 1999). Such learning enables the visual system to encode and retain the stable structure of the environment, so that subsequent search can be guided efficiently by memory.

In contextual cuing studies, target location is learned relative to a particular context. But what, precisely, constitutes a *context* for the guidance of visual search? Recent studies have wrestled with this thorny issue, examining how the spatial relationship between target location and a repeated search context is encoded (Brockmole et al., 2006; Jiang & Wagner, 2004; Olson & Chun, 2002). Studies using arrays of simple stimuli have suggested that target position is learned relative to the local context in which the target appears, with participants most sensitive to the spatial configuration of the objects closest to the target (Brady & Chun, 2007; Olson & Chun, 2002). Olson and Chun (2002) divided “repeated” arrays into two spatial regions. In one of the regions, the configuration of contextual elements was held constant across trials. In the other, the positions of the contextual elements varied randomly. Contextual cuing was found only when the target appeared within the static half of the display. That is, repetition of the local context around the target was sufficient to generate contextual cuing (despite variability in the other half of the display), and local variability around the target completely blocked contextual cuing (despite the fact that the other half of the display perfectly predicted target location). In addition, Brady and Chun (2007) found that repetition limited to the two distractors closest to the target was sufficient to produce contextual cuing equivalent to that produced by repetition of the entire array of 11 distractors. These results suggest that target location is learned relative to the local context immediately adjacent to that object.

In a similar vein, Jiang and Wanger (2004) found evidence that target location is learned relative to the individual locations of contextual objects rather than relative to the global configuration of the search display. During a learning phase, a particular target location was repeated within two different search contexts. In a transfer block, half of the distractors from each context were combined into a hybrid display. All of the individual distractor locations had been associated with the target location during the learning phase, but the global configuration of distractors in the hybrid display matched neither of the two contexts in which target position had been learned. Strikingly, contextual cuing was undi-

minished in this condition, suggesting little or no role for global configuration in contextual cuing.<sup>1</sup>

In contrast with these findings, Brockmole et al. (2006) reported evidence of global dominance in a contextual cuing paradigm using images of real-world environments. The scenes used by Brockmole et al. contained a small search target (a letter). One scene item was presented multiple times during a learning phase. This item contained the target on a table within a library scene. In a transfer block, the local context (table) and global context (everything else in the library scene) were decoupled. No contextual cuing was observed if the table on which the target appeared was presented in a different scene. However, contextual cuing was preserved, and was undiminished, if the target appeared at the same location in the library scene but on a different table. Changing the global context eliminated contextual cuing, whereas changing the local context had no effect on contextual cuing. In a second experiment, repeating only the local context (i.e., presenting the table in a series of different scenes) led to minimal facilitation of search, whereas repeating only the global context led to significantly greater learning. Brockmole et al. reasoned that, unlike randomly generated arrays (Brady & Chun, 2007; Jiang & Wagner, 2004; Olson & Chun, 2002), natural scene stimuli have an inherent spatial structure that can be used to organize information about the locations of individual objects in memory. In addition, rapid recognition of the scene based on its global properties could provide an efficient means to retrieve information about the location of the target object. In general, then, the Brockmole et al. results suggest that the positions of objects in natural scenes are learned, primarily, relative to the global structure of the entire scene.

In sum, there are two seemingly contradictory views regarding the nature of spatial information that guides search through scenes. On the one hand, there is strong evidence that local spatial cues guide search through repeated search arrays (Brady & Chun, 2007; Jiang & Wagner, 2004; Olson & Chun, 2002); on the other hand, there is evidence that global spatial cues guide search through repeated natural scenes (Becker & Rasmussen, 2008; Brockmole et al., 2006). Resolving the apparent contradiction between evidence of local dominance and global dominance in contextual cuing is complicated by the many stimulus differences between the respective experiments. More generally, theorizing about the role of global and local context in the learning of target position is complicated by the fact that *global* and *local* are not clearly defined, especially within real-world scenes. What is to be considered a “local” context in a natural scene? Do the contents of a desk drawer constitute a local context? Do the “items by the window” constitute a local context? And is a global context “everything else in the scene,” as operationalized by Brockmole et al. (2006), or is a global context limited to objects at a particular spatial scale or governed by the geometric structure of the bounding features of the scene (such as walls and floor)?

<sup>1</sup> Jiang and Wagner (2004) also found that contextual cuing survives translation and expansion of the array, which could be interpreted as evidence for a role of abstract, configural representations in contextual cuing. But their method did not allow them to determine whether this effect was driven by the global configuration of the display or by the local information adjacent to the target.

If one has entered an office and is searching for a desk, then the context for search would be the office. If one is now standing in front of the desk and searching for a stapler on the desktop, the desk would constitute the bounding context for the search. One could even search the stapler to find the label listing the type of staples it uses, in which case the stapler itself could be considered the context for search. Because the functional context for search can be instantiated at multiple spatial scales, it is unlikely that search is dominated, in general, by either global or local structure. Such distinctions are possible only in relationship to a particular example of search behavior. Brockmole et al. (2006) also acknowledged this difficulty in defining global context. For example, they noted that although a particular object in the display was considered a local cue, that same object could be considered part of the global representation within another search context.

In addition, each search behavior is likely to depend on memory representations at a particular spatial scale optimized for the task at hand. When searching the stapler for a label, search will be guided by knowledge of the typical location of labels on staplers (underneath the base) or memory for the location of the label on one's own stapler. One certainly would not learn the location of the stapler label relative to the large-scale structure of the office (or the office building) in which it appears, especially as the location of the stapler in the office might change. It is therefore unlikely that one's memory for the large-scale structure of the office would play any significant role at all in finding the stapler label, except to the extent that knowledge of the office can guide attention initially toward the stapler itself. Similarly, when searching the refrigerator for milk, it is likely that search is guided by one's memory for the location of the milk relative to the spatial structure of the refrigerator and is not governed by memory for the location of the milk relative to the large-scale structure of the kitchen (or house) in which the refrigerator appears. Memory for the spatial structure of the kitchen allows one to find the refrigerator efficiently. And memory for the spatial structure of the refrigerator allows one to find the milk efficiently. At each stage, successful search at one spatial scale leads to the retrieval of contextual information at a finer spatial scale until the ultimate target object is found.

### The Present Study

In the present study, we sought to test the hypothesis, developed above, that memory for a particular search context is nested within memory for the larger scene in which it appears. In addition, we sought to reconcile, at least in part, the conflicting results regarding the nature of the contextual learning in visual search (Brockmole et al., 2006; Jiang & Wagner, 2004; Olson & Chun, 2002). We began with the assumption that object position often is learned relative to a functional subregion of an environment. That is, the position of the milk in the refrigerator will be learned relative to the spatial structure imposed by a subregion of the kitchen, the refrigerator, and not relative to the global structure of the kitchen or house. This assumption is broadly consistent with the results of Olson and Chun (2002) and Brady and Chun (2007), showing that target location can be learned relative to a subregion of a search array. We refrain from using the term *local context* to refer a subregion of a scene precisely because the global–local distinction is dependent on the scale of the search and is difficult to define within real-world environments. To probe this type of learning, the

experiments in the present study confined visual search to well-defined subregions of depicted scenes.

We further hypothesized that the expression of this learning will depend on recognition of the larger scene itself, consistent with an important role for scene recognition in contextual cuing (Brockmole et al., 2006; Brockmole & Henderson, 2006a, 2006b). That is, to efficiently retrieve a structured memory representation of the layout of the refrigerator and the position of the milk, one first must establish that one is inside the appropriate kitchen. A representation of the broader scene must be instantiated to efficiently gain access to information about the structure of subregions within that scene, whether the subregion is a kitchen within a house, a refrigerator within the kitchen, or a shelf within the refrigerator. This hypothesis is based on the assumption that memory for complex environments is organized hierarchically (McNamara, Hardy, & Hirtle, 1989), with information about functional subregions of an environment bound to locations within a representation of the environment at a larger spatial scale.

Evidence for this type of contextual structure in natural scenes comes from a recent study by Hollingworth (2007). Participants' access to memory for the visual form of an individual object in a scene was superior when that object was presented at the same position within the scene at which it had been viewed originally (see also Hollingworth, 2006). However, this position specificity was observed only when the original scene context was presented at test. If the test object was presented against a blank background, no same-position advantage was obtained, suggesting that access to memory for the location of an object was dependent on instantiating the original scene context in which it appeared. Similarly, we propose that access to memory for the spatial structure of a subregion of a scene will be dependent on presenting that subregion within the original scene context.

In our approach, then, we do not treat “global” and “local” spatial structure as dichotomous, and we do not propose that the learning of target location in natural scenes is dominated by information at a particular spatial scale. Instead, we propose that contextual learning at one spatial scale is nested within a representation of the larger scene, and that access to spatial context in visual search is dependent on activating a representation of the larger environment in which that information was learned. To test this hypothesis, we developed a contextual cuing paradigm in which participants searched through well-defined subregions of a scene. Each *search region* was embedded within a *scene background* in a manner that allowed us to manipulate the relationship between the search region and the scene in which it appeared.

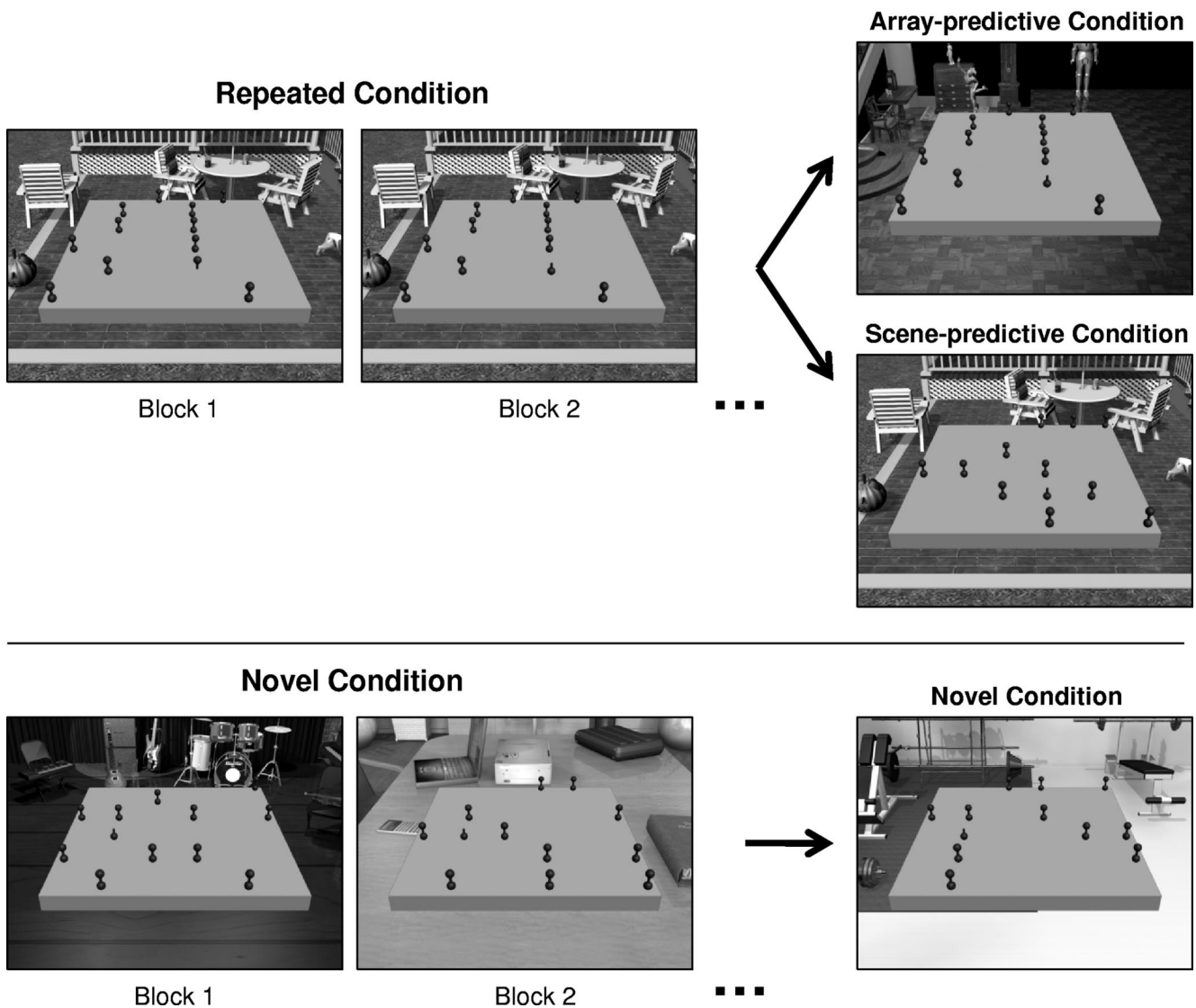
### Experiment 1

The basic paradigm is illustrated in Figure 1. On each trial, participants viewed a stimulus composed of a central search region (green “table”) surrounded by a scene background depicting a real-world environment. The search region contained an array of 12 “dumbbell” objects. Participants searched for a target that was missing one globe and reported which globe, top or bottom, was missing (target type was assigned randomly on each trial). The target always appeared within the search region. Search response time (RT) was the dependent measure.

The experiment consisted of two sessions. In the training session, participants completed 24 blocks of 16 search trials. Eight of

## Training Session

## Transfer Session



*Figure 1.* Stimulus manipulations in Experiment 1. In the repeated condition of the training session, each target location was paired with a consistent search array configuration and scene background in every block. In the novel condition of the training session, each target location was paired, randomly, with a search array and scene background in each block. In the transfer session, array-predictive trials paired a learned array and target location with a randomly chosen scene background. Scene-predictive trials paired a learned scene background and target location with a randomly generated search array.

the trials in each block were in the *repeated* condition. For repeated stimuli, a particular target location was consistently paired with a particular search array configuration and with one of eight scene backgrounds. As illustrated in Figure 1, the entire stimulus (search array and scene background) was repeated in each block, except for the target type, which randomly varied, forcing participants to find the target on every trial. The repeated condition was contrasted with a *novel* condition. In each block, a particular target location was presented within a randomly generated search array and within a scene background randomly chosen from a set of

eight scene backgrounds. We expected to find significant contextual cuing in the training session, with faster RTs in the repeated condition than in the novel condition as participants learned the relationship between repeated stimuli and target location.

The training session was followed, without pause, by a transfer session of six blocks, in which the relationship between the search region and scene background was manipulated. Each participant completed one of two transfer conditions. In the *scene-predictive* transfer condition, scene backgrounds and target locations from the repeated condition were paired in each transfer block with a novel



configuration of distractors in the search array. Scene background still predicted target location, but the configuration of the search array did not. In the *array-predictive* transfer condition, search arrays and target locations from the repeated condition were paired, randomly in each block, with scene backgrounds that had appeared in the novel condition during training. The configuration of the search array still predicted target location, but the scene background did not. Each predictive transfer condition was compared with a novel condition in which neither the scene background nor the search array configuration predicted target location.

Because search was always limited to the central search region, we expected that the specific location of the target object in repeated stimuli would be learned relative to the spatial structure of that search region and not relative to the large-scale structure of the scene background (an assumption confirmed in Experiments 2 and 3). Thus, we expected that in the scene-predictive transfer condition, contextual cuing would be significantly reduced or eliminated, as the search array context no longer specified target location. We predicted a similar result in the array-predictive transfer block, but for a different reason. In the array-predictive transfer block, the search array context was preserved. However, if the expression of array learning depends on instantiating the original scene in which a particular array context was learned, the presentation of a learned array within a different scene background should impair retrieval of the array context and should reduce or eliminate contextual cuing in this condition as well. In summary, we predicted that neither the scene background nor the array would be sufficient for the transfer of contextual cuing: Disrupting the array context should reduce or eliminate contextual cuing because the position of the target was learned relative to the search array; disrupting the scene background should also reduce or eliminate contextual cuing because the retrieval of the relevant array context would be impaired.

These predictions contrast with predictions generated from evidence that contextual cuing is dominated by either local (Olson & Chun, 2002) or global information (Brockmole et al., 2006). If target location is coded relative to local elements independently of the larger scale identity of the scene, then contextual cuing should be unreduced in the array-predictive transfer condition and should be eliminated in the scene-predictive transfer condition. If target location is coded relative to the global elements of the scene, then contextual cuing should be unreduced in the scene-predictive transfer condition and should be eliminated in the array-predictive transfer condition.

## Method

**Participants.** Participants in all of the experiments were members of the University of Iowa community between the ages of 18 and 30 years. They received course credit or pay. All participants reported normal or corrected-to-normal vision. Each participant completed only one experiment. Forty-eight participants completed Experiment 1, with 24 in each of the two transfer conditions.

**Stimuli.** Search stimuli comprised two segregated regions: search arrays and scene backgrounds. Search arrays consisted of 12 red “dumbbell” objects arranged randomly in a  $6 \times 6$  grid on the surface of a green “table.” These arrays were then superimposed into the center of a set of scene backgrounds depicting

real-world environments. All stimuli were generated using 3DStudio Max software. There was a total of 16 different scene backgrounds, each with highly distinct visual and semantic content (e.g., “farm,” “living room,” “picnic table,” “playground”). Background scenes contained depth cues consistent with the search displays to provide good visual integration of the scene background and search region.

Search arrays of relatively simple, meaningless objects were used so that the search array itself would be unlikely to contribute to identification of the scene background. If we had used common, real-world objects in the search arrays, then it is possible that the recognition of those objects would play a role in the recognition of a particular scene background paired with those objects. By using arrays of identical, meaningless objects in every scene, we could change the pairing of scene backgrounds and search arrays in the transfer session without significantly influencing the identity of the scene background. In Figure 1, note that the patio scene in the training and transfer sessions is clearly the same patio, despite the fact that the search array has been changed.

Search arrays were randomly generated, with the constraint that the target could not appear in the back row (because of the possibility that a target in the back row could intersect directly with the background image). The green table subtended  $14.1^\circ$  vertically. Horizontally, it subtended  $16.2^\circ$  at the bottom and  $11.7^\circ$  at the top. The size of the search elements (“dumbbells”) ranged from  $0.65^\circ$  (width)  $\times$   $1.45^\circ$  (height) in the front row of the search array to  $0.43^\circ \times 1.05^\circ$  in the back row of the search array. The target was a dumbbell missing either the top or the bottom globe. The entire stimulus image (scene background and search array) subtended  $22.7^\circ$  (width)  $\times$   $17.0^\circ$  (height).

**Apparatus.** The stimuli were displayed at a resolution of  $800 \times 600$  pixels  $\times$  24-bit color on a 17-in. CRT monitor with a refresh rate of 100 Hz. Responses were collected using a serial button box. The presentation of stimuli and collection of responses were controlled by E-Prime software running on a PC-compatible computer. The room was dimly illuminated by a low-intensity light source. A forehead rest maintained a viewing distance of 80 cm.

**Procedure.** On each trial, participants first saw a gray screen with the word “ready?” in the center. When ready, the participant pressed a button on the button box to initiate the trial. There was a blank (gray screen) image presented for 250 ms, followed by the search stimulus. The search stimulus remained visible until response. Participants pressed the *left* button on the response box if the top globe was missing and the *right* button if the bottom globe was missing. They were instructed to respond as quickly as possible without making a significant number of errors. Incorrect responses were followed by the text “incorrect” for 500 ms. There was a 500-ms delay between trials.

**Training session.** Participants completed 24 blocks of 16 trials in this phase of the experiment. Half of the trials in each block were in the repeated condition and half in the novel condition. In the repeated condition, there were eight target locations, and each target location was used once in every block of the training session. In addition, each target location was uniquely paired in every block with a particular array of search distractors and with a particular scene background. Thus, both the search array and the scene background predicted target location. In the novel condition, there was also a set of eight target locations, and each location was used once in every block. However, within a

block, each of these target locations was paired with a randomly generated array of distractors and with a scene background drawn randomly from the set of eight scene backgrounds not used in the repeated condition. Thus, neither the scene background nor the search array configuration predicted target location.

The two conditions were equated for the number of target locations (8) used throughout the training session. The 16 target locations were chosen randomly without replacement from the 30 possible target locations in the  $6 \times 6$  grid (excluding the back row). To ensure that target locations were counterbalanced across conditions, the assignment of target positions to the repeated and novel conditions for one participant was reversed for the next participant. In addition, we controlled the number of times each individual scene background was presented in each condition (once in each block). The assignment of the 16 scene backgrounds to the two conditions was determined randomly for each participant.

Within each block, trials from the repeated and novel conditions were randomly intermixed.

**Transfer session.** After the 24 blocks of training, participants completed one of two transfer sessions. The transfer session consisted of six blocks of 16 trials. There was no division between the training session and the transfer session. In the array-predictive transfer condition, the search arrays from the repeated training condition remained predictive of target location, but they were paired, randomly, with different scene backgrounds. In the scene-predictive transfer condition, the scene backgrounds from the repeated training condition remained predictive of the target location, but they were paired with randomly generated arrays of search distractors.

Trials in the scene-predictive transfer blocks were divided evenly between repeated trials, which maintained the relationship between scene background and target location, and novel trials. In the repeated condition, the target locations used in the repeated trials of the training session continued to be paired with the same scene backgrounds with which they had been paired during training. However, the configuration of distractor elements in the search array was randomly generated on each trial. Scene background continued to predict target location, but the search array did not. In the novel condition of the scene-predictive transfer block, the eight target locations from the novel condition of the training session continued to be paired with a scene background randomly chosen from the set of eight used in the training session and with a randomly generated configuration of search distractors.

Trials in the array-predictive transfer blocks were evenly divided between repeated trials, which maintained the relationship between the search array and target location, and novel trials. In the repeated condition, the target locations used in the repeated trials of the training session continued to be paired with the same distractor arrays as in training. However, the scene background in which the repeated array appeared was randomly selected in each block from the eight scene backgrounds that had appeared in the novel condition during the training session (these scene backgrounds had not been consistently associated with any particular target location). The search array continued to predict target location, but the scene background did not. In the novel condition of the array-predictive transfer block, the eight target locations from the novel condition of the training session were paired with a randomly generated configuration of search distractors and with a

scene background randomly chosen from the set of eight that had appeared in the repeated condition of the training session.<sup>2</sup>

## Results

Overall, subjects were highly accurate at performing this task, with mean accuracy of 96.6% correct on novel trials and 96.7% correct on repeated trials. RT analyses were limited to correct responses. Response times less than 100 ms or greater than 5,000 ms were eliminated as outliers, and were no more than 0.1% of the data in any experiment.

**Training data.** To assess the results from the training phase of our experiment, we used a new statistical technique for the analysis of contextual cuing data: nonlinear mixed-effects modeling (NLME; Pinheiro & Bates, 2000). NLME is a statistical analysis that is most often used in fields such as pharmacology (Davidian & Giltinan, 2003), but its growing popularity is evidenced by adoption in other fields, such as biology (Wu & Ding, 1998), medicine (Pauler & Finkelstein, 2002), and experimental psychology (Young, Clark, Goffus, & Hoane, 2009). This analysis produces two main types of results: parameters that fit a function for the data, such as a power function or exponential function, computed for the various conditions in the experiment (e.g., “repeated trials” and “novel trials”); and statistical tests over these parameters to judge goodness of fit and differences between the experimental conditions.

NLME has several major advantages for analyzing repeated search data over more commonly used repeated measures analysis of variance (ANOVA). We describe these briefly before reporting the results of Experiment 1. The advantages of using NLME for analyzing repeated search data are similar to the advantages discussed in Young et al. (2009) for modeling Morris water maze data in rats. Chiefly, the standard ANOVA approach is insensitive to the fact that time (in our case, “search block”) is a continuous, rather than a categorical, variable, which results in decreased power. Furthermore, the two main effects in an ANOVA are not typically of primary interest. If the search type factor (repeated vs. novel) is significant, this indicates that repeated contexts are easier than novel contexts irrespective of the number of trials the participant has completed, thus removing the learning component of contextual cuing. If the search block factor is significant, this simply indicates that in some blocks, the participant searched faster than in other blocks (without respect to the evolution of learning as the experiment progressed). Thus, the main term of interest from a repeated measures design is the interaction of search type (novel or repeated) with trial block. But, even in the case that an interaction is found, a series of follow-up tests is needed to identify the specific blocks in which the difference was significant. Most important, ANOVA does not provide a means to identify the functions characterizing learning in each condition. To describe these functions, a popular secondary method is to fit functions for

<sup>2</sup>This last feature of the design could be problematic if participants learned the location of the target relative to the scene background, as this might influence search performance in the novel condition. Search RTs could have been elevated because the target was not presented in the position that had been associated with the background scene during training. However, the results of Experiment 1, reported below, rule this out as a concern.

each subject and perform statistical analyses on the parameters (e.g., Chun & Jiang, 2003). Using this type of method, however, removes a large amount of within-subject variability, reducing the likelihood of finding a significant difference and potentially mischaracterizing the data set.

NLME solves each of these issues. The key question in a contextual cuing experiment is whether there are differences in learning rate as a function of repetition. The main analysis should concern the slope of the function that describes search latency in each condition over the course of the experiment. Thus, it is appropriate to treat the “search block” variable as continuous rather than categorical, and the NLME treatment of search block as continuous will generally result in a large increase in power over more traditional methods. In addition, NLME is sensitive to the within-subject variability that is lost by fitting parameters for subjects individually. The main output of the NLME technique is the parameters for a general function that describes the learning data in each of the conditions of the experiment. These parameters are fit by the model converging on best estimates for the subject and population parameters simultaneously.

We fit our training session data with a two-parameter power function in the form of  $RT = ix^s$  (see Chun & Jiang, 2003). The  $i$  parameter describes the intercept of the function, the  $s$  parameter describes the slope of the function, and  $x$  corresponds to search block. A statistically significant difference in the intercept parameter corresponds to a difference in baseline difficulty between the two conditions irrespective of search block. A difference in slope corresponds to progressively faster searches through some types of displays as a function of search block.

Although it is possible to manually define the intercept parameter to be equivalent between repetition conditions as an a priori assumption, NLME fits were computed with both slope and intercept parameters as fixed effects to test the validity of that assumption. A reliable fit to a power function was obtained for both the novel and repeated trials in the training session. For novel trials, a value of 1027 was obtained for the intercept parameter and  $-.055$  for the slope parameter. For repeated trials, a value of 1050 was obtained for the intercept parameter and  $-.082$  for the slope parameter. The functions computed with these parameters, along with the average RT data, are shown in Figure 2. Although the intercept parameter fits were highly reliable,  $F(1, 2205) = 758.4, p < .001$ , there was no significant difference between these parameters,  $F(1, 2205) = 1.0, p = .31$ , indicating that repeated displays were not generally easier than novel displays. For the slope parameter, there was also a significant fit to the data,  $F(1, 2205) = 34.7, p < .001$ . Critically, there was also a significant difference in the slope parameter between conditions,  $F(1, 2205) = 15.62, p < .001$ . Thus, there was a reliable effect of search repetition, as indicated by the difference in slope parameters that defined the best-fit power functions for the learning data. Participants were progressively faster at searching through repeated displays relative to novel displays as the experiment progressed. Over the last 12 training blocks, the average size of this benefit was 49 ms.

The training session was identical for the two transfer groups. However, to confirm that both groups produced a contextual cuing effect, we conducted separate NLME analyses for the two and found them to be highly similar. When analyzed separately, neither group showed a significant difference in the intercept parameter: array-predictive group,  $F < 1$ ; scene-predictive group,  $F(1,$

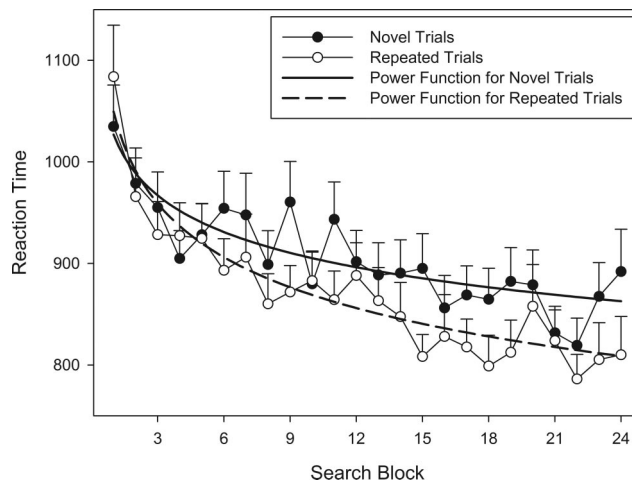


Figure 2. Mean search times for repeated (open circles) and novel (filled circles) trials in the Experiment 1 training session. The best-fitting power functions (dashed and solid lines, respectively) obtained by nonlinear mixed effects modeling are superimposed onto the training conditions (details in text). Error bars represent standard errors of the means.

1101) = 1.38,  $p = .24$ ; but both showed highly significant slope parameter differences: array-predictive group,  $F(1, 1101) = 7.27, p = .007$ ; scene-predictive group,  $F(1, 1101) = 7.89, p = .005$ .

**Transfer data.** As can be seen in Figure 3, contextual cuing was significantly reduced in the transfer sessions of both transfer conditions. Changing either the scene background or the search array that was paired with a particular target location significantly reduced any effect of contextual learning obtained during the training session. Thus, neither the search array context nor the scene background was sufficient for the robust transfer of contextual cuing.

To analyze the transfer data, we returned to the more common repeated measures ANOVA as the variables were categorical (i.e., training vs. transfer). Specifically, we compared the repetition effect in the six blocks of the transfer session with the repetition effect in the last 12 blocks of the training session (in which the contextual cuing effect had reached asymptote). We first conducted an omnibus ANOVA with repetition (repeated, novel) and session (training, transfer) as within-subject factors and transfer group (scene predictive, array predictive) as a between-subjects factor. There was a no main effect of group, nor any reliable interaction between group and the other factors, suggesting that the two groups produced similar data patterns across the factors of interest.

Critically, there was a reliable interaction between repetition (repeated, novel) and session (training, transfer),  $F(1, 46) = 12.3, p = .001$ . A robust contextual cuing effect was observed in the last 12 blocks of training,  $F(1, 46) = 13.7, p < .001$ , with mean RT in the repeated condition 49 ms faster than mean RT in the novel condition. However, there was not a reliable contextual cuing effect in the transfer session,  $F < 1$ . The numerical trend in the transfer session was toward a slightly negative contextual cuing effect, with mean RT in the repeated condition 3 ms slower than mean RT in the novel condition (a contextual cuing effect of  $-3$  ms). The experiment had sufficient power to detect a contextual

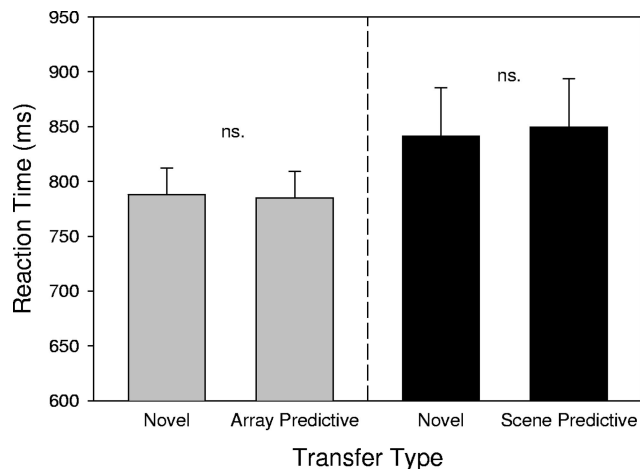


Figure 3. Mean search times in the array-predictive (left) and scene-predictive (right) transfer sessions of Experiment 1. Error bars are 95% within-subject confidence intervals based on the error term of the repetition effect for each transfer group.

cueing effect of 24 ms in the transfer block. Thus, if a contextual cueing effect of the magnitude observed during training had been present in the transfer session, the experiment had sufficient power to detect it.

In addition to the omnibus ANOVA, we analyzed the data from the two transfer groups separately. For the array-predictive group, there was a reliable interaction between repetition and session,  $F(1, 23) = 5.82, p = .02$ , with a 61-ms contextual cueing effect during the last 12 blocks of training,  $F(1, 23) = 8.49, p = .008$ , but only a 3-ms contextual cueing effect in the transfer session,  $F < 1$ . Contextual cueing was reliably disrupted by changing the scene backgrounds associated with previously learned arrays. The 95% confidence interval for the contextual cueing effect in the transfer session was  $3 \text{ ms} \pm 25 \text{ ms}$ . Although we cannot state conclusively that contextual cueing was eliminated entirely in the transfer session, we can be confident that any contextual cueing effect was no larger than 28 ms.

For the scene-predictive group, there was likewise a reliable interaction between repetition and session,  $F(1, 23) = 7.06, p = .01$ , with a 37-ms contextual cueing effect during the last 12 blocks of training,  $F(1, 23) = 5.36, p = .03$ , but a  $-8$ -ms contextual cueing effect in the transfer session,  $F < 1$ . Changing the local array associated with a repeated target location significantly disrupted contextual cueing. The 95% confidence interval for the contextual cueing effect in the transfer session was  $-8 \text{ ms} \pm 44 \text{ ms}$ . Thus, we can be confident that any contextual cueing effect in the transfer session was less than 36 ms.

One concern with the design of the array-predictive transfer session is that for the novel condition of that transfer session, target locations were paired with scene backgrounds that had appeared in the repeated condition of the training session. It is possible that participants learned the relationship between these scene backgrounds and particular target locations during training. When these scene backgrounds were paired with new target locations in the transfer session, participants may have directed attention to the original, learned location, increasing RT. This would result in artificially high RTs in the novel condition of the transfer session

compared with the repeated condition, potentially producing an effect that would mimic a contextual cueing result. This possibility does not constitute a major concern, however. First, no such pattern of results was observed in the array-predictive transfer session. Participants were no slower in the novel condition than in the repeated condition, and so there was no evidence that the prior relationship between the scene background and a target location generated higher RTs in the novel condition. This provides strong evidence that the target location was not learned relative to the scene background. The results of Experiment 2 provide converging evidence that target location is not learned relative to the scene background in this paradigm.

## Discussion

In Experiment 1, we found that contextual cueing developed over the course of a training session as search arrays and background scenes were consistently paired with target locations. However, in a transfer session, neither the array context nor the scene background was sufficient for the transfer of contextual cueing. Contextual cueing was significantly reduced (and numerically eliminated) by a change to the array context and by a change to the scene background. The stimuli were constructed so that search was limited to a well-defined subregion of a scene, and we expected that target location would be learned relative to the spatial configuration of the search region. The disruption of contextual cueing with change to the search array context is consistent with that assumption. Despite the fact that the target was presented within the same scene as it had originally appeared (i.e., at the same location within the same patio scene), changing the immediate search context around the target blocked the transfer of contextual cueing. This result is broadly consistent with the results of Olson and Chun (2002) showing that a change in “local” context impairs contextual cueing. In addition, the disruption of contextual cueing in the scene-predictive transfer block contrasts with the result of Brockmole et al. (2006), in which changes to the “local” context in scenes had no effect on transfer. We discuss the likely causes of the difference between our result and Brockmole et al.’s in the General Discussion.

Second, we hypothesized that when search array context was preserved and reliably predicted target location, expression of array learning would be dependent on instantiating the original scene background in which the search array had been learned. Consistent with this hypothesis, a predictive search array context generated no observable contextual cueing when paired with a new scene background. This result is broadly consistent with the proposal that contextual cueing is dependent on scene recognition (Brockmole et al., 2006; Brockmole & Henderson, 2006a). It is inconsistent with the finding that global and distal disruptions have no detrimental effect on contextual cueing (Jiang & Wagner, 2004; Olson & Chun, 2002) and therefore eliminates the hypothesis that only “local” information is functional in contextual cueing.

Together, these results suggest that memory for object location can be learned relative to functional subregions of an environment (such as memory for the spatial structure of one’s refrigerator), and that memory for the subregion is nested within a representation of the larger scene (kitchen), with efficient access to the spatial structure of the former requiring instantiation of the latter.



## Experiment 2

Although hierarchical environmental representation is a plausible cause of the Experiment 1 results, it is not the only possible cause. An alternative explanation for the lack of transfer in the array-predictive condition is that target position was learned, at least to some extent, relative to the large-scale structure of the scene. That is, contextual cuing was disrupted not because instantiating the scene background was necessary for retrieval of the predictive search region context, but because changing the scene background directly disrupted memory for the location of the target object. This possibility is limited by the results of the novel trials in the array-predictive transfer condition. In these trials, scene backgrounds that had been consistently paired with target locations during training were paired with new target locations in the transfer session. Yet, there was no observable cost in the transfer session, as would have been expected if memory for the position of the target relative to a particular scene background had guided attention during search.

Nevertheless, in Experiment 2 we tested directly whether a consistent pairing between scene background and target location would be sufficient to support contextual cuing. Participants completed a single session of 24 blocks of 16 trials. Each target location in the repeated condition was consistently paired with a particular scene background. However, the search array configuration of distractors was randomly generated on each trial. The repeated condition was compared with a novel condition that was structured in the same manner as in Experiment 1.

## Method

**Participants.** Twenty-four participants completed Experiment 2.

**Stimuli and apparatus.** Stimuli were generated in the same manner as in Experiment 1. The apparatus was the same as in Experiment 1.

**Procedure.** The procedure was the same as in Experiment 1, except that during the session, only the scene background predicted target location on repeated trials. Specifically, eight target locations were used for the repeated condition. Each of these locations was paired in every block with one of the eight scene backgrounds used for this condition. However, the configuration of distractors in the search array was randomly generated on every trial. For the novel trials, there were also eight target locations. In each block, each location was paired with a scene background randomly selected from the set of eight scene backgrounds used for this condition. In addition, the configuration of distractors in the search array was randomly generated on every trial. Thus, in the novel condition, neither array context nor scene background predicted target location. In the repeated condition, only the scene background predicted target location. There was no transfer session.

## Results and Discussion

RT analyses were limited to correct responses. Accuracy on the search task was high, with mean accuracy of 97.2% correct on novel trials and 96.8% correct on repeated trials.

Learning data were again analyzed by fitting power functions (of the form  $RT = ix^c$ ) using NLME. These best-fitting curves,

along with the average learning data, are plotted in Figure 4. Significant fits were obtained for both novel and repeated displays. Novel trials were fit by a function with an intercept parameter of 1084 and a slope parameter of  $-.07$ ; repeated trials were fit by a function with an intercept parameter of 1075 and a slope parameter of  $-.07$ . Once again, the intercept parameters did not significantly differ between conditions. Critically, there was no significant slope difference between the two repetition conditions,  $F(1, 1101) < 1$ . That is, there was no evidence that participants could use the consistent pairing of target locations and scene backgrounds to facilitate search in the repeated condition. During the last 12 blocks of the session, the mean contextual cuing effect was 5 ms. We compared the magnitude of this effect with the magnitude of the contextual cuing effect observed in the last 12 blocks of the Experiment 1 training session. Consistent with the presence of contextual cuing in the latter but not the former, there was a near-significant interaction between repetition (repeated, novel) and experiment,  $F(1, 70) = 3.80, p = .055$ .

These results indicate that scene backgrounds were insufficient to produce reliable cuing benefits in the absence of a predictive search array context. Furthermore, they discount the possibility that the loss of cuing in the array-predictive transfer condition of Experiment 1 was due to disruption of memory for the location of the target object relative to the scene background. If such learning occurred, it produced no observable contextual effect on search in Experiment 2. This conclusion is further strengthened by the finding that, in Experiment 1, pairing potentially predictive scene backgrounds with new target locations generated no cost in search RT. Thus, the data suggest that for the present stimuli, target location was not coded relative to the scene background (or that such coding did not contribute significantly to contextual cuing).

## Experiment 3

The results of Experiments 1 and 2 indicate that target locations were learned relative to the subregion of the scene over which

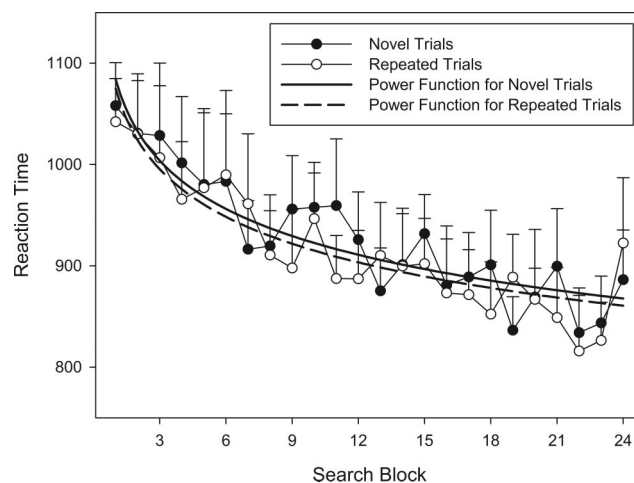


Figure 4. Mean search times for repeated (open circles) and novel (filled circles) trials in Experiment 2. The best-fitting power functions (dashed and solid lines, respectively) obtained by nonlinear mixed effects modeling are superimposed onto the training conditions. Error bars represent standard errors of the means.

search occurred (the search region), but that expression of this learning was dependent on instantiating the original scene context in which the search array appeared. In Experiment 3, we examined whether consistent search array context would be sufficient, by itself, to support learning and contextual cuing in the absence of a consistent scene background during training. The method was complementary with that of Experiment 2: On repeated trials, the search array configuration predicted target location, but the scene background did not.

## Method

**Participants.** Thirty participants completed Experiment 3.

**Stimuli and apparatus.** Stimuli were generated in the same manner as in Experiments 1 and 2. The apparatus was the same as in Experiments 1 and 2.

**Procedure.** The procedure was the same as in previous experiments, except that only the search array context predicted target location on repeated trials. Eight target locations were used in every block with one of eight distractor arrays. However, the background scene was randomly chosen in every block from one of the 16 scene backgrounds. In the novel condition, each of the eight locations was paired with a randomly generated distractor array and a scene background randomly selected from the set of 16 scene backgrounds. Thus, in the novel condition, neither array context nor scene background predicted target location. In the repeated condition, only the array context predicted target location.

## Results and Discussion

RT analyses were limited to correct responses. Mean accuracy was 96.8% correct on novel trials and 96.6% correct on repeated trials.

Statistical analyses of these data were again conducted by fitting power functions (of the form  $RT = ix^n$ ) using NLME. Significant fits were obtained for both novel and repeated displays. These best-fitting curves, along with the average learning data, are plotted in Figure 5. Novel trials were fit by a function with an intercept parameter of 1105 and a slope parameter of  $-.07$ ; repeated trials were fit by a function with an intercept parameter of 1133 and a slope parameter of  $-.10$ . Although the intercept parameters for both conditions were a statistically reliable fit, they did not differ between conditions,  $F(1, 1377) < 1$ . The slope parameter was significantly fit to the data, and there was a significant slope difference between the two repetition conditions,  $F(1, 1377) = 8.16, p < .005$ , demonstrating significant contextual cuing. In the last 12 blocks of the session, the mean contextual cuing effect was 31 ms. The presence of contextual cuing is consistent with our assumption that target location was learned relative to the configuration of the search region. The 31-ms contextual cuing effect was numerically smaller than the contextual cuing effect of 49 ms observed during the training session of Experiment 1, in which both the array context and the scene background predicted target location. However, a between-experiments comparison did not produce a reliable difference in the magnitude of contextual cuing,  $F < 1$ . Overall, the difference in learning rates between the two search conditions (novel vs. repeated) was 23% smaller in Experiment 3 than in Experiment 1.

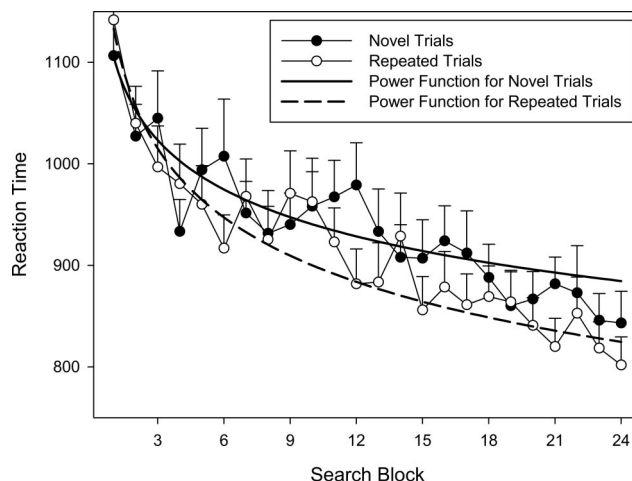


Figure 5. Mean search times for repeated (open circles) and novel (filled circles) trials in Experiment 3. The best-fitting power functions (dashed and solid lines, respectively) obtained by nonlinear mixed effects modeling are superimposed onto the training conditions. Error bars represent standard errors of the means.

In Experiment 3, search through repeated arrays was faster than search through novel arrays, even though there was no consistent pairing between repeated search arrays and scene backgrounds. Thus, a consistent scene background is not a necessary condition for contextual cuing; the structure of search regions can be learned in a manner that tolerates scene variability, consistent with the results observed in Experiment 2 of Brockmole et al. (2006), as well as in Olson and Chun (2002). This result generates an intriguing contrast with the results of Experiment 1. For static search contexts that are always paired with a single scene environment, as in Experiment 1, the expression of array learning is strongly scene-dependent. However, for repeated search contexts that appear in multiple scenes, as in Experiment 3, the search array can still be learned, but not in the strongly hierarchical and scene-specific manner observed in Experiment 1. Presumably, memory for search arrays learned within multiple scenes would be more likely to transfer to novel scenes than memory for search arrays learned within a single scene. We are currently testing this possibility.

## General Discussion

In the present study, we sought to reconcile conflicting evidence suggesting that contextual learning in visual search is dominated either by local information immediately surrounding the target (Brady & Chun, 2007; Olson & Chun, 2002) or by global properties of the environment, such as the identity of the scene in which the target appears (Brockmole et al., 2006). This issue strikes to the heart of how spatially organized environmental representations are structured in memory and how such memory representations interact with visual attention to enable efficient search. We hypothesized that the specific location of an object is often learned relative to a functional subregion of environment (such as learning the location of the milk relative to the structure of the refrigerator rather than relative to the larger scale structure of the kitchen in which the refrigerator appears), an assumption broadly consistent

with evidence that the facilitation of visual search can be driven by repetition of a local context around the target location (Brady & Chun, 2007; Olson & Chun, 2002). However, we further hypothesized that the expression of such learning within real-world scenes would be dependent on instantiating the original scene context in which learning occurred, consistent with evidence that contextual cuing in meaningful, real-world environments is dependent on scene recognition (Brockmole et al., 2006; Brockmole & Henderson, 2006a).

Search stimuli were created in which the relationship between a discrete search region and the scene background could be manipulated. Participants showed reliable learning of repeated search arrays, as indicated by contextual cuing, and contextual cuing was observed even when there was no consistent pairing between repeated search arrays and scene backgrounds. However, when a particular repeated array had been consistently associated with a particular scene background, the expression of array learning was contingent on presenting the search array within the original scene background; there was no reliable transfer of contextual cuing when a repeated array associated with one scene background was presented in a different scene background.

These data suggest that knowledge for the spatial structure of a functional subregion of an environment is nested within a representation of the larger scene. For example, although the position of the milk is unlikely to be remembered relative to the large-scale spatial structure of the kitchen, retrieval of one's knowledge of the structure of the refrigerator depends on first recognizing the kitchen. Similar hierarchical structure has been found in studies of human spatial memory. For example, when recalling the locations of objects in an environment or in a map, individual locations are grouped into spatial regions, and access to those locations is ordered in a region-by-region manner (Hirtle & Jonides, 1985; McNamara, 1986; McNamara et al., 1989; Taylor & Tversky, 1992). From this type of evidence, McNamara et al. (1989) argued that internal representations of complex, large-scale environments consist of multiple maps at different spatial scales, with maps at finer spatial scales hierarchically nested within maps at coarser spatial scales. Similar hierarchical structure has been proposed in the literature on spatial memory and navigation in rats, with access to memory for individual food locations organized by functional subregions of the larger environment (Macuda & Roberts, 1995). It is well established that this type of spatial–environmental learning is strongly dependent on the hippocampus in rats (Bostock, Muller, & Kubie, 1991; Morris, Garrud, Rawlins, & O'Keefe, 1982; O'Keefe & Speakman, 1987). Considering that contextual cuing in humans is impaired by medial temporal damage (Chun & Phelps, 1999), it is plausible that for the purpose of visual search, medial temporal regions play a role in establishing the association between memory for a functional subregion of an environment (such as the spatial structure of the refrigerator) and memory for the larger scene in which it appears.

The present results diverge somewhat from those of contextual cuing studies using traditional visual search arrays of randomly placed, simple stimuli (Brady & Chun, 2007; Jiang & Wagner, 2004; Olson & Chun, 2002). As discussed above, contextual cuing under these circumstances can be explained almost entirely by local, distractor-to-target associations. For example, Brady and Chun (2007) were able to simulate much of the extant contextual cuing data using a model that imple-

mented learning as a collection of individual target–distractor associations, weighted so that learning was driven primarily by the distractors immediately adjacent to the target. It is certainly possible that such an approach could account for the learning of target position relative to the central search array in the present experiments. That is, the learning of the specific target location in repeated arrays might very well have depended on local target–distractor associations within the search array. However, such localist accounts cannot easily accommodate the effects of scene background observed in Experiment 1, as they have no reliable mechanism for generating an influence of large-scale scene structure or meaning. To some extent, this discrepancy is likely to reflect the fact that random arrays of objects simply do not instantiate the large-scale spatial structure of natural scenes (i.e., that natural scenes consist of coherent structural elements such as walls, counters, and floors) or the fact that natural scenes are easily identifiable and discriminable at the level of scene category (e.g., “kitchen”) and scene exemplar (e.g., “Sue's kitchen”). Given that the ultimate goal of contextual cuing research is to understand the guidance of attention during search through familiar, natural scenes, theories and models of contextual cuing will need to be extended to reflect the fact that real-world search almost always occurs in structured, meaningful environments.

In addition, the present results diverge from those found in studies examining contextual cuing in natural scenes. Brockmole et al. (2006) observed that preserving the “global” elements of a scene (i.e., everything in the scene except a table on which the target appeared) was sufficient for the transfer of contextual cuing, whereas we found that preserving the scene background was not sufficient for the transfer of contextual cuing if the configuration of the search array was changed. In addition, Brockmole et al. found that repeating only the “global” context generated significant contextual cuing, whereas we found that repeating only the scene background was insufficient to generate contextual cuing.

These different results almost certainly derive from differences in the structure of the scene stimuli and the nature of the search task in the two studies. In Brockmole et al. (2006), the target letter could appear anywhere within the scene. The relevant search region therefore constituted the entire scene stimulus. In contrast, our stimuli were designed so that search would be limited to a well-defined subregion of the scene. Thus, it is not particularly surprising that the learning of target location in the present experiments was driven strongly by the configuration of the search region, as participants had no reason to search for the target anywhere else in the scene. One can view our method and the Brockmole et al. method simply as two types of search through natural environments. The Brockmole et al. method represents a case in which participants have no information about the plausible regions of a scene in which the target might appear, and thus they must search the entire scene. Target location is learned relative to the functional context for search: the entire scene. The present method represents a case in which participants know the subregion of the scene that is relevant for search, similar to knowing that the milk is almost certainly located in the refrigerator. Target location was likewise learned relative to the functional context for search: the central search array. Thus, the difference between the two studies can be viewed as a difference in the functional context for

search. As discussed in the beginning of this article, this should not be interpreted necessarily as a dichotomy between “global” and “local” contexts, as search contexts can be instantiated at multiple spatial scales.

Despite these differences, our results converge with those of Brockmole et al. (2006) in suggesting a critical role for scene recognition in contextual cuing. The present study extends our understanding of this relationship by demonstrating that scene identity can control contextual cuing even when the scene background itself is not immediately task relevant. Participants could have performed the search task without attending to or encoding any information from the scene backgrounds at all, because the target never appeared within the scene background region. Yet, a change to the pairing between repeated search arrays and scene backgrounds disrupted contextual cuing.

This is not to say that efficient visual search will always depend on scene identification. For example, when we learn the spatial structure of the inside of a car or the arrangement of objects in a backpack, we do so despite the fact that these contexts are moveable and can appear within many different scene environments. Such scene-independent spatial learning, which supports search within contexts that appear in variable locations, may require the explicit decoupling of a search context from a particular environment. This was the case in Experiment 3, in which target locations were learned without the consistent pairing of scene background and repeated array. However, the default strategy when presented with a consistently paired search context and scene background is likely to be one of hierarchical coding. When a search array has been consistently associated with a single scene, the expression of array learning appears to be scene-dependent.

Ultimately, the debate regarding the importance of either global or local features in guiding visual search may have depended too strongly on the assumption that “global” and “local” contextual structures can be distinguished and are separable within a particular example of visual search. In the present experiments, we provided strong perceptual cues to segregate search arrays from scene backgrounds. Yet, even under conditions that should have been optimized for observing context segregation in search, we observed that neither the search array nor the scene background was sufficient for the transfer of contextual cuing when the two had been consistently paired. Thus, the influence of scene knowledge on visual search should not be considered a case of cue competition (in which one source of information, such as “global” or “local” spatial structure, comes to dominate attentional guidance). Rather, the influence of scene knowledge on visual search should be considered a case of cue integration, with memory for the structure of a functional subregion of a scene nested within a representation of the larger environment.

## References

- Becker, M. W., & Rasmussen, I. P. (2008). Guidance of attention to objects and locations by long-term memory of natural scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1325–1338.
- Bostock, E., Muller, R. U., & Kubie, J. L. (1991). Experience-dependent modifications of hippocampal place cell firing. *Hippocampus*, 193–205.
- Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 798–815.
- Brockmole, J. R., Castelano, M. S., & Henderson, J. M. (2006). Contextual cuing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 699–706.
- Brockmole, J. R., & Henderson, J. M. (2006a). Recognition and attention guidance during contextual cuing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 1177–1187.
- Brockmole, J. R., & Henderson, J. M. (2006b). Using real-world scenes as contextual cues for search. *Visual Cognition*, *13*, 99–108.
- Castelano, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 753–763.
- Chen, X., & Zelinsky, G. J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research*, *46*, 4118–4133.
- 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. H. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 224–234.
- Chun, M. M., & Phelps, E. A. (1999). Memory deficits for implicit contextual information in amnesic patients with hippocampal damage. *Nature Neuroscience*, *2*, 844–847.
- Chun, M. M., & Turk-Browne, N. B. (2008). Associative learning mechanisms in vision. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 209–246). New York: Oxford University Press.
- Davidian, M., & Giltinan, D. M. (2003). Nonlinear models for repeated measurement data: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, *4*, 387–419.
- Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science*, *17*, 973–980.
- Henderson, J. M., Brockmole, J. R., Castelano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford, England: Elsevier.
- Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228.
- Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, *13*, 208–217.
- Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 58–69.
- Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 31–47.
- Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*, *17*, 273–291.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Jiang, Y. H., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—Configuration or individual locations? *Perception & Psychophysics*, *66*, 454–463.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*, 3559–3565.



- Macuda, T., & Roberts, W. A. (1995). Further evidence for hierarchical chunking in rat spatial memory. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 20–32.
- McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*, 87–121.
- McNamara, T. P., Hardy, J. K., & Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 211–227.
- Morris, R. G. M., Garrud, P., Rawlins, J. N. P., & O'Keefe, J. (1982, June 24). Place navigation impaired in rats with hippocampal lesions. *Nature*, *297*, 681–683.
- Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research*, *46*, 614–621.
- Neider, M. B., & Zelinsky, G. J. (2008). Exploring set size effects in scenes: Identifying the objects of search. *Visual Cognition*, *16*, 1–10.
- O'Keefe, J., & Speakman, A. (1987). Single unit activity in the rat hippocampus during a spatial memory task. *Experimental Brain Research*, *68*, 1–27.
- Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, *9*, 273–302.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123.
- Pauler, D., & Finkelstein, D. M. (2002). Predicting time to prostate cancer recurrence based on joint models for non-linear longitudinal biomarkers and event time outcomes. *Statistics in Medicine*, *21*, 3897–3911.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer-Verlag.
- Pomplun, M. (2006). Saccadic selectivity in complex visual search displays. *Vision Research*, *46*, 1886–1900.
- Taylor, H. A., & Tversky, B. (1992). Descriptions and depictions of environments. *Memory & Cognition*, *20*, 483–496.
- Torralba, A., Oliva, A., Castelhana, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*, 766–786.
- Wu, H., & Ding, A. A. (1998). Population HIV-1 dynamics in vivo: Applicable models and inferential tools for virological data from AIDS clinical trials. *Biometrics*, *55*, 410–418.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.
- Young, M. E., Clark, M. H., Goffus, A., & Hoane, M. R. (2009). Mixed effects modeling of Morris water maze data: Advantages and cautionary notes. *Learning and Motivation*, *40*, 160–177.

Received March 12, 2009

Revision received August 31, 2009

Accepted September 11, 2009 ■

### Online First Publication

APA-published journal articles are now available Online First in the PsycARTICLES database. Electronic versions of journal articles will be accessible prior to the print publication, expediting access to the latest peer-reviewed research.

All PsycARTICLES institutional customers, individual APA PsycNET® database package subscribers, and individual journal subscribers may now search these records as an added benefit. Online First Publication (OFP) records can be released within as little as 30 days of acceptance and transfer into production, and are marked to indicate the posting status, allowing researchers to quickly and easily discover the latest literature. OFP articles will be the version of record; the articles have gone through the full production cycle except for assignment to an issue and pagination. After a journal issue's print publication, OFP records will be replaced with the final published article to reflect the final status and bibliographic information.