

## Spatial Constraints on Learning in Visual Search: Modeling Contextual Cuing

Timothy F. Brady  
Massachusetts Institute of Technology

Marvin M. Chun  
Yale University

Predictive visual context facilitates visual search, a benefit termed *contextual cuing* (M. M. Chun & Y. Jiang, 1998). In the original task, search arrays were repeated across blocks such that the spatial configuration (context) of all of the distractors in a display predicted an embedded target location. The authors modeled existing results using a connectionist architecture and then designed new behavioral experiments to test the model's assumptions. The modeling and behavioral results indicate that learning may be restricted to the local context even when the entire configuration is predictive of target location. Local learning constrains how much guidance is produced by contextual cuing. The modeling and new data also demonstrate that local learning requires that the local context maintain its location in the overall global context.

*Keywords:* visual search, contextual cuing, computational model, statistical learning

The visual world is full of an overwhelming amount of information, and much of this information is not immediately useful to individuals at any given moment. In a given scene, there are dozens or even hundreds of objects and events competing for visual awareness, and yet observers may be looking for only one specific object—or may be trying to ignore all of them. This huge amount of largely irrelevant information would be impossible to process and deal with if one had no way to sort and choose what to attend. Attentional mechanisms help individuals to focus on a specific set of objects and events (Chun & Wolfe, 2001; Kanwisher & Wojciulik, 2000; Pashler, 1998; Treisman & Gelade, 1980) and therefore prioritize the scene and attend to only relevant stimuli. Researchers have identified a number of visual cues that attract attention. Visual search tasks, for example, have suggested that some objects enter attention because they simply pop out from a scene if they possess unique or salient features (Egeth, Jonides, & Wall, 1972; Treisman & Gelade, 1980; Wolfe, 1994; Yantis, 1998) or if they abruptly onset (Yantis & Jonides, 1984). Although such bottom-up cues are very useful when noticing a bright red stop sign on a suburban street, they are less useful when it comes to noticing a red traffic light in the middle of Times Square in Manhattan, New York City. Yet people do somehow reliably detect traffic lights even amidst hundreds of other stimuli that are just as salient.

Beyond visual salience, an important cue for attention may be contextual information that informs which objects should appear in a scene and where. Objects and events almost always occur in a rich visual context (Biederman, 1972; Chun, 2000). This context

tends to be highly predictable, because one's visual experience is not based on a random sample of objects; it is highly structured and repetitive, a fact that the visual system takes advantage of (E. J. Gibson, 1963, 1991; J. J. Gibson, 1966). Thus, objects are associated with the context in which they are usually found, helping to facilitate recognition of objects in their normal context (Bar & Ullman, 1996; Boyce, Pollatsek, & Rayner, 1989; Chun & Jiang, 1999; Mackworth & Morandi, 1967; but see Hollingworth & Henderson, 1998). Likewise, context information constrains the positions of objects within scenes (Biederman, Mezzanotte, & Rabinowitz, 1982; Hollingworth, 2006; Palmer, 1975). This helps cut down on the massive information overload because it provides constraints on the range of possible objects that can be expected to occur in a particular context (e.g., visual experience suggests that a long tan object in someone's living room is probably a couch, not a lion). Contextual information allows the visual system to benefit from the fact that the visual world is, in general, stable—what people see in a particular context one day is likely to still be there the next day.

The visual system's sensitivity to context has also been demonstrated using a more highly controlled visual search task, in a paradigm known as *contextual cuing* (Chun, 2000). The original contextual cuing paradigm showed that observers implicitly learn the global configuration of targets in visual search tasks and that this context can serve to cue the target location and facilitate search performance in subsequent encounters (Chun & Jiang, 1998). In the generic contextual cuing experiment, observers perform a visual search for targets appearing among distractor stimuli arrayed in invariant (predictive) or variable (nonpredictive) spatial configurations, randomly intermixed within blocks (see Figure 1 for a sample display). Predictive configurations are repeated across blocks throughout the entire session, and targets appear in consistent locations within these configurations. Sensitivity to global context leads to faster target search performance in these repeated configurations compared with the nonpredictive configurations that are regenerated in each block. This learning is implicit, as

---

Timothy F. Brady, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; Marvin M. Chun, Department of Psychology, Yale University.

Correspondence concerning this article should be addressed to Timothy F. Brady, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 46-4078, 77 Massachusetts Avenue, Cambridge, MA 02139. E-mail: tbrady@mit.edu

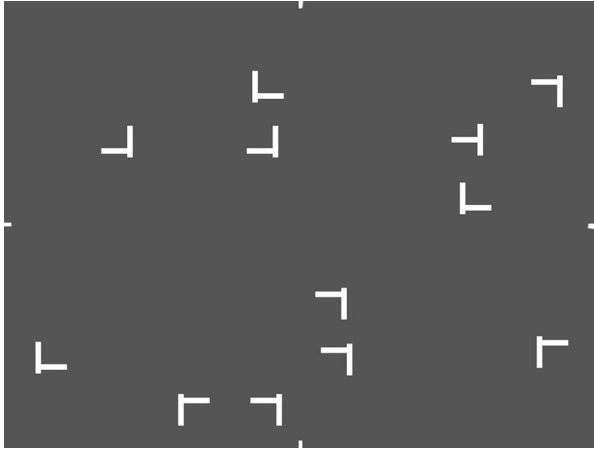


Figure 1. Example contextual cuing display. In Chun and Jiang (1998), the entire display would be repeated from block to block. In Olson and Chun (2002), only one side of the screen (in this example, the right side containing the target in the short-range predictive condition and the left side in the long-range predictive condition) would be repeated.

observers are at chance in predicting which quadrant the target should appear in when given both predictive and nonpredictive configurations in a memory test (Chun & Jiang, 2003).

Newer work has suggested that observers are not equally sensitive to the entire invariant configuration, however. Olson and Chun (2002, Experiment 1) demonstrated that attention is cued more strongly by the configuration neighboring the target location than by the configuration on the opposite side of the screen. In this experiment, the invariant configuration occupied only one half of the screen in predictive trials, with the other half of the screen variable between repetitions. The division between variable and invariant configuration was imperceptible because no visual boundaries existed between the two halves of the screen. Thus, the difference between signal and noise was defined statistically over trials. The target appeared either within the invariant half of the screen (short-range predictive condition) or within the variable half (long-range predictive condition). Observers showed significant contextual cuing in the short-range predictive condition but not in the long-range predictive condition, indicating that they were not equally sensitive to the entire configuration. Olson and Chun (2002) explained their result as a difficulty in learning through intervening noise. However, a more general account is possible. This article will introduce a model and behavioral experiments that demonstrate how learning is restricted to the local area around the target, constraining how much statistical information is encoded from the displays.

### Connectionist Model

We modeled Chun and Jiang's (1998, Experiment 1) and Olson and Chun's (2002, Experiment 1) contextual cuing experiments using a two-layer neural network to help understand what is learned in these contextual cuing tasks and what the computational limitations on this learning might be. This model formalizes the assumptions underlying contextual cuing, including the units it operates over, the spatial extent of the learning, and the role of

attentional selection. It also attempts to unite a wide range of findings on contextual cuing into a simple model to make the mechanisms of contextual cuing more explicit.

This model tested contextual learning only. A number of other computational models have been designed to describe how the visual input is processed to support target search among distractors (Bundesen, 1990; Cave, 1999; Humphreys & Muller, 1993; Itti & Koch, 2000; Logan, 1996; Wolfe, 1994), but our model operates independently of the specific details of visual processing. The only input necessary for our model is a representation that specifies which locations contain search items and which do not—in other words, configuration information.<sup>1</sup> In this sense, our model can be applied to almost any other model of visual search to implement contextual cuing.

The network receives the displays as input in matrix form and then processes them on the basis of its past experience. The network then outputs values for each location in the screen matrix, representing what it estimates to be the likelihood the target is present there. After each target is found, the network adjusts its weights so that the next time it sees the same input, it is more likely to rank the correct target location highly, which in turn would prioritize the deployment of attention and eye movements (Peterson & Kramer, 2001). Thus, the network learns in the same style that human observers must be: associating the target location with its surrounding context to help facilitate future searches.

The network's input nodes are given their values from a screen matrix with a value of 0 assigned if no object is present at the location and a value of 1 assigned if there is an object present. The output nodes are treated as an activation map, with one output node per screen location. Each of these output nodes is given a value between 0 and 1 indicating how likely the network believes it is to be the target location. For each trial, the appropriate screen is given as an input to the network, and its output is calculated using a feed-forward mechanism based on two sets of previously assigned weights and a sigmoid transfer function. The first set of weights is the model's learned weights, and the second is a fixed set of weights built into the model. The fixed weights are used to implement spatial constraints in the model and will be discussed further below. Additionally, we included a bottom-up activation term, which increases the value of any output node when its corresponding input node is active. This makes the network more closely resemble human behavior by having it search primarily those locations in the screen matrix that actually contain items (an assumption that is implicit in telling people to identify the target).

Mathematically, this means that for any output node,  $y_i$ , where  $W1$  is the learned weight matrix,  $W2$  is the fixed weight matrix,  $\beta$  is the bottom-up activation component, and  $x_j$  is an input node,  $y_i$ 's value is derived using the following:

$$y_i = 1/(1 + e^{-s_i})$$

$$s_i = \sum_k [x_k W1(x_k, y_i) W2(x_k, y_i)] + (\beta x_i).$$

Once this output has been calculated, the network "looks" at the locations that produced the highest activations in its output layer, continuing in decreasing order until the target is found. The

<sup>1</sup> The assumption that contextual cuing operates over configuration information will be examined in the Modeling Results section.

learned weights are then updated using the delta rule, with the error term ( $\epsilon$ ) for a given output node being its distance from the correct output (1 if it corresponded to the target location and 0 in all other locations), before the network moves onto the next trial. So, where  $W1(x_k, y_i)_t$  is the value of the learned weight between input node  $x_k$  and output node  $y_i$  at time  $t$ ,  $\mu$  is the learning rate, and  $m$  is the momentum, the new value is computed using the following:

$$\delta(x_k, y_i)_{t+1} = (\mu\epsilon)(y_i)(x_k) + [m\delta(x_k, y_i)_t]$$

$$W1(x_k, y_i)_{t+1} = W1(x_k, y_i)_t + \delta(x_k, y_i)_{t+1}.$$

Momentum ( $m$ ) is set to 0.95, learning rate ( $\mu$ ) is set to 0.001, with bottom-up activation ( $\beta$ ) at 0.1.

Over many trials, this results in each input node activating only those output nodes that have been target locations while the input node was active. This means that when a display is repeated, the nodes with the highest activations tend to be those that have been paired as target locations with a large subset of the distractor locations. After learning, this converges to the correct target location. The network's search time is computed as a function of how many locations in the matrix it looks at before finding the target (e.g., the number of locations in the output layer that have a higher activation than the target location), as an approximation to human response time. This search time does not include any of the fixed costs of visual search (segregation of items from background, response selection, motor planning and execution, etc.).

The contextual cuing paradigm is represented as closely as possible, with the network going through the same number of learning blocks and trials used in the original behavioral experiments. Each time the network is run, it emulates 1 observer, starting with randomly assigned learned weights and going through the entire experiment. Reported network outputs are the result of averaging together as many runs of the network as observers were used in the equivalent human experiment.

### Spatial Constraints on Network Connectivity

The fixed weights were used to impose spatial constraints on the learning of the model. This was necessary in order to make it compatible with the Olson and Chun (2002, Experiment 1) data showing that human observers are more sensitive to a short-range than a long-range context. This higher sensitivity to local context in search behavior indicates that any connectionist implementation of contextual cuing must not have full connectivity between the input and output layers of the network. This constraint is necessary because a fully connected network cannot encode spatial information—all of the output nodes have equal access to all of the input nodes, which means that with the proper adjusting of the network's weights, it is as easy for the network to learn an association between the target location and a set of distant stimuli as an association between the target location and nearby stimuli.

Our particular account of spatial constraints is based on the idea of an attentional spotlight that surrounds the target when it is localized, allowing the spatial relationships within the spotlight to be encoded. The strength of the learning for a given distractor is based on how far from the center of the spotlight it is located.

In this account of spatial constraints, no arbitrary borders are hard coded into the display. Likewise, the constraints are continuous, such that two nodes are neither connected or not connected,

but instead the strength of their connection is modulated by the distance between the two nodes. This account is the result of considering the computational limitations inherent in the task and the need to explain both the original Olson and Chun (2002, Experiment 1) limitation in learning far from the target and the possibility that learning at a distance may sometimes be possible (see the Modeling Results section and Olson & Chun, 2002, Experiment 3).

Mathematically, the spatial constraints are represented using an exponential function:

$$W2(x_k, y_i) = e^{-d(k,i)},$$

where  $d$  is distance and  $d(k, i)$  represents the number of matrix locations separating the  $k$ th input node from the  $i$ th output node. Figure 2 provides a graphical view of the fixed weights from all of the input nodes to one particular output node (in this case, the location of the target).

In sum, spatial configurations are made explicit in the model via the addition of a separate set of fixed weights that implement our account of spatial constraints in contextual cuing.

### Model Results

We initially tested the network on the original experiment by Chun and Jiang (1998, Experiment 1). We used an  $8 \times 6$  screen matrix, with 12 objects per screen (11 distractors and 1 target) and ran 30 blocks of 24 trials each, just as in the original experiment. Half of the trials in a given block were predictive trials, and half were nonpredictive. The model data very closely approximated the human data. At first, the network was no better at guessing the location of the target than random looking throughout the input, but as it was exposed to the repeated configurations, it learned to associate the configuration with the target locations. As indicated by Figure 3, the network shows significant cuing ( $p < .01$  in all epochs except one) after a short learning period.

The model also captures another element of human performance in contextual cuing tasks by showing an improvement in response time for nonpredictive displays. It demonstrates this improvement because only half of the 48 possible screen locations are actually target locations in this experiment. The model automatically extracts these target location probabilities over time, using them to limit its search to only potential target locations—an effect often

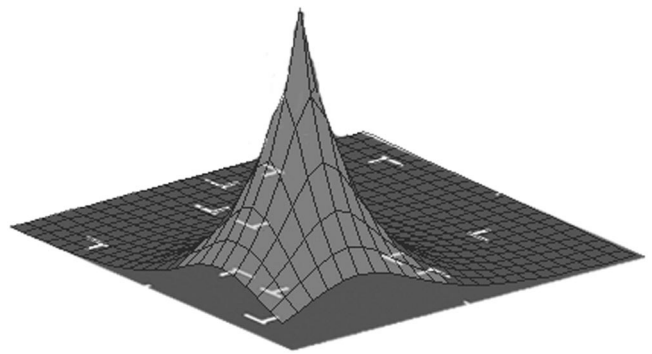


Figure 2. Modulation weights in the model, for an output node corresponding to the target location in this picture.

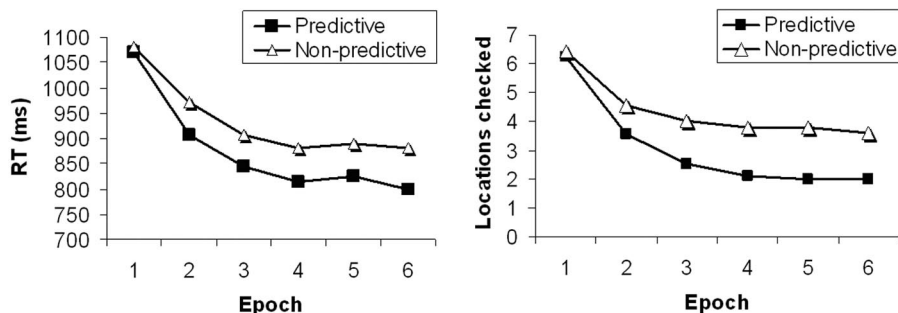


Figure 3. Data from the original contextual cuing experiment (as in Chun & Jiang, 1998): Human data are on the left, and model data are on the right. The model data are an average of 32 runs of the model, each representing 1 observer. The human data are based on 16 participants. RT = response time. Data in the left panel are from “Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention,” by M. M. Chun and Y. Jiang, 1998, *Cognitive Psychology*, 36, p. 36. Copyright 1998 by Elsevier. Adapted with permission.

observed in human observers (e.g., Miller, 1988). The model does this automatically as a byproduct of the way it learns from context.

Directly comparing the model and human learning curves in a nonarbitrary way is not possible, because the model does not include the fixed costs of search (e.g., segregation of items from background, response selection, motor planning and execution). To get some idea of the degree of this target location learning in the model, however, we can take advantage of the fact that it results in a benefit of approximately the same magnitude as contextual cuing. Because human observers showed a contextual cuing benefit of between 60 and 80 ms in Chun and Jiang’s (1998) experiment and because the model showed a benefit of approximately two locations checked, each location the model is checking corresponds to a 30- 40-ms increase in response time, comparable with prior estimates of the rate of attentional shifts in tasks like this (Wolfe, 1994).

Human observers exhibit overall practice effects of nearly 200 ms (Chun & Jiang, 1998), suggesting that if human observers are similarly extracting target location probabilities, more than 120–140 ms of the overall practice effect is likely due to factors other than learning target location probabilities (e.g., better segregation of the displays, improved shape discrimination of the target shapes

from distractors, motor learning, etc.). These other practice effects are outside the scope of our model, which attempts to account for contextual cuing only, rather than visual search more generally. Thus, we will focus on the differences between old and new conditions.

Because learning in the model is based on the delta rule, the amount of change in the weights is directly proportional to the amount of error the model makes in predicting the location of the target. At the beginning, when this error is very high, the weight changes are very large and make a large difference in the predictions of the model. As the amount of error decreases, the amount of change in the weights in each trial decreases. The model thus displays an exponential learning curve that is very typical of neural networks in general, especially those using a form of gradient descent like the delta rule (Haykin, 1998). This exponential shape is very similar to the learning curves displayed by contextual cuing in previous experiments, which were found to fit both an exponential and a power function well (Chun & Jiang, 2003).

In response to this learning, the magnitude of contextual cuing increases roughly logarithmically in both the model and human observers. As shown in Figure 4, the best fit curve for both contextual cuing in the model and contextual cuing in human

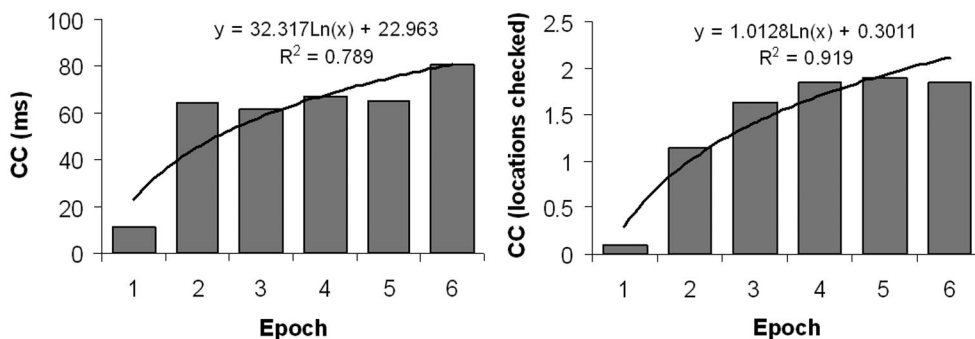


Figure 4. Magnitude of contextual cuing (from Chun & Jiang, 1998): Human data are on the left, and model data are on the right. CC = contextual cuing. Data in the left panel are from “Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention,” by M. M. Chun and Y. Jiang, 1998, *Cognitive Psychology*, 36, p. 36. Copyright 1998 by Elsevier. Adapted with permission.

observers is logarithmic, with  $r^2 = .79$  for human observers and  $r^2 = .92$  for the model, both accounting for a very large amount of the variance. By comparison, the best fit linear model gives  $r^2 = .61$  for the human data and  $r^2 = .73$  for the model data, accounting for less of the variance. In both the model and human observers, therefore, the greatest magnitude of contextual learning takes place early in the experiment.

After assessing how well the model fit the general trend of contextual cuing, we next ran the model on the Olson and Chun (2002, Experiment 1) experiment that prompted the inclusion of spatial constraints. As in the behavioral experiment, four conditions were used: short-range predictive contexts, long-range predictive contexts, global-predictive contexts, and completely non-predictive contexts. To generate the displays, we divided screens in half vertically, with the invariant context being on the same side as the target in the short-range predictive condition, on the opposite side than the target in the long-range predictive condition, and on both sides in the global-predictive condition. There were no visible boundaries to distinguish the two sides. We used a  $12 \times 8$  screen matrix and ran 20 blocks of 32 trials each (8 trials of each condition per block).

The results shown in Figure 5 indicate that the model's data very closely approximated the human data, in terms of both contextual cuing and the limits imposed by the spatial separation of the predictive context from the target. Long-range cuing was negligible, whereas short-range cuing was moderate and comparable with the global-predictive condition. Cuing in the model was significant for the short-range and global-predictive conditions in all but Epoch 1 (all  $ps < .01$ ) and was insignificant in all epochs for the long-range predictive condition (all  $ps > .10$ ). Thus, our model correctly captured the results of Olson and Chun (2002, Experiment 1).

It is important to note that others have previously modeled contextual learning using the selective attention for identification model of visual search (Backhaus, Heinke, & Humphreys, 2005; Heinke & Humphreys, 2003). To do so, they used a Hopfield-like memory with asymmetrical weights. These weights represented the observed covariation between two spatial locations and biased the network's attentional selection layer toward settling on the

target location. This model succeeded in capturing both the basic contextual cuing effect and the effects of set size on contextual cuing, the first model to do so. However, their model did not address the question of spatial constraints on learning, a major, novel focus of our model that helps explain a much wider range of results, as detailed later in the Modeling Results section of this article. In theory, similar spatial constraints could be implemented in their model, and we suggest that, given the similarity of the two models, they would come to many of the same conclusions we have in this article.

Once we had established that our network could accurately model basic contextual cuing, we tested its predictions with new behavioral experiments. In particular, we wanted to test how robust learning could be with limited predictive information embedded in nonpredictive noise. No parameters of the model described above were changed to generate the specific predictions reported below. Four experiments tested the model's predictions with human performance. Afterwards, we used the same model to simulate a large body of results from the broader contextual cuing literature. New insights are gained from such simulations, such as answers to why attentional guidance from contextual cuing is not perfect.

### Experiment 1

We first tested whether it is possible to decrease the amount of predictive information available to benefit observers in their search for the target and still obtain contextual cuing. Rather than have one half of the distractors be predictive as in Olson and Chun (2002, Experiment 1), we limited the predictive information to only the distractors in the target quadrant. Thus, rather than the 8 distractors out of 15 that were predictive of target location in Olson and Chun (2002, Experiment 1), only 2 distractors out of 11 were predictive of the target location in this experiment (see Figure 6). This represents an extremely noisy and limited amount of predictive information for observers to learn from each of the repeated displays. If observers are limited by spatial constraints like those posited by our model, this task should be no more difficult than the Olson and Chun (2002, Experiment 1) task. However, if observers

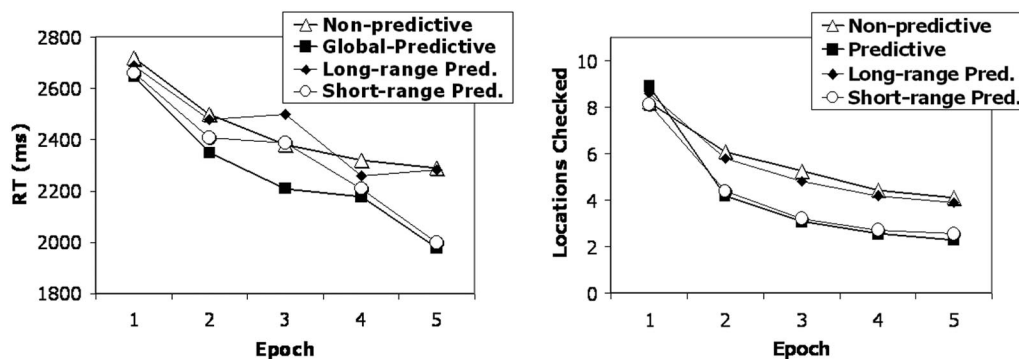


Figure 5. Data from Olson and Chun (2002), using contexts in which only half the screen was invariant and predictive. Human data are on the left, and model data are on the right. RT = response time; Pred. = predictive. Data in the left panel are from "Perceptual Constraints on Implicit Learning of Spatial Context," by I. R. Olson and M. M. Chun, 2002, *Visual Cognition*, 9, p. 281. Copyright 2002 by Psychology Press ([www.psypress.co.uk/journals.asp](http://www.psypress.co.uk/journals.asp)). Adapted with permission.

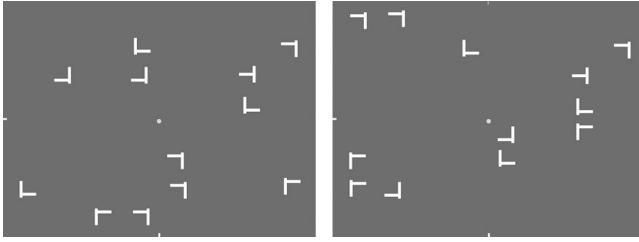


Figure 6. Example displays from Experiment 1. In one block, the quadrant-predictive trial would look like the display on the left; in the next block, it might look like the one on the right. The target quadrant (top right) is the only invariant area.

do not have such spatial constraints, learning should be extraordinarily difficult in this experiment: There is far more noise than signal in each display, and even side by side the displays barely resemble each other (see Figure 6).

### Method

**Observers.** Fourteen members of the Yale University community participated in this experiment in partial fulfillment of an introduction to psychology course requirement or as paid volunteers (\$10/hour for approximately 1 hr). All observers reported normal or corrected-to-normal visual acuity and normal color vision. None of the observers were aware of the purpose of this study.

**Stimuli.** The target was a *T*-shaped stimulus rotated 90° to either the left or the right. The distractor stimuli were *L* shapes that were rotated to appear randomly in any of four possible orientations (0°, 90°, 180°, or 270°). The *L* shapes were modified such that the two line segments that would meet at a 90° angle in a typical *L* were offset by 33%, making the *L*s more similar to the target stimuli. This made the search task more difficult for observers (Duncan & Humphreys, 1989). On each trial, there was exactly one target, and observers pressed the arrow key (left or right) corresponding to the direction the bottom of the *T* was pointed. Each display contained 12 items (11 distractors and 1 target), which were distributed across an 8 × 8 matrix of locations not visible to the observer. The locations were constrained such that an equal number (3) of items appeared in each of the four quadrants of display. In addition, targets were prevented from appearing directly near the center of the display or at the corners of the display, creating a bias for them to appear within a donut-shape set of locations around fixation. This was done in order to control for eccentricity and to avoid excessively easy (at fixation) or excessively hard (in the corners) target locations. As a result, targets were more likely to be aligned with local context items within the same quadrant than across quadrants. The target quadrant contained 1 target item and 2 other distractor items. The distractor quadrants always contained 3 distractors.

All experiments were performed on a Macintosh computer using MATLAB (The MathWorks, Natick, MA) with the Psychophysics Toolbox extensions (Brainard, 1997). The viewing distance was approximately 55 cm, although observers were unrestrained. Stimuli were about 1.8 × 1.8 degrees in visual angle, and were aligned in an invisible 8 × 8 grid that took up the entire screen. The

position of each item was jiggled so as to prevent collinearities with other stimuli. The background was set to a dark gray, and the stimuli were always white.

**Design and procedure.** The two main variables were configuration (nonpredictive vs. quadrant predictive) and block. The quadrant-predictive trials consisted of 16 randomly generated configurations in which the target quadrant configuration remained constant throughout the experiment. In other words, the target location and the layout of distractors within the quadrant were fixed for each target quadrant configuration. On the quadrant-predictive trials, the identity of the distractor items in the target's quadrant were preserved, but the target direction was changed from repetition to repetition so that the correct response did not correlate with any of the configurations it appeared in. The distractor configuration in the other three quadrants was newly generated in each block. The nonpredictive trials consisted of 16 different configurations, which were entirely newly generated for each block to serve as a control. Distractor configurations were chosen randomly from all possible locations, including target locations from other trials.

The experiment was conducted in two phases: training (20 blocks) and test (10 blocks). During the training trials, observers were exposed only to the 16 quadrant-predictive configurations. During the test trials, these same quadrant-predictive trials configurations were interspersed with 16 nonpredictive trials that shared the same target locations, in order to control for effects based on target location probabilities. Each session consisted of 20 blocks of 16 trials during training and 10 blocks of 32 trials during test, for a total of 640 trials per observer. Each configuration appeared exactly once per block.

The observer pressed the space bar to begin each block. Each trial started with a small fixation dot appearing in the middle of the screen. After a pause of 500 ms, the stimuli appeared on the screen. They remained there until either the observer pressed a button or 8 s had passed. Observers were directed to press the arrow key corresponding to the target as soon as they had identified the target. The response caused the screen to clear, with the word *Error* appearing if the observer responded incorrectly. A beep was also sounded if the observer did not respond correctly or if he or she took more than 8 s to respond. The response was followed by a pause of 1,500 ms and then the next trial was automatically started. Following each block of trials, the observers were given a break of at least 5 s but were allowed to wait as long as they wished before hitting the space bar to begin the next block. At the conclusion of the experiment, observers were presented with 32 configurations with the target replaced with a distractor and were asked to guess which quadrant the target should appear in (Chun & Jiang, 2003). Half of these trials were entirely nonpredictive, and the other half were the quadrant-predictive trials used throughout the experiment.

The experiment began with instructions followed by a practice block of 16 trials to allow the observers to adjust to the task. Observers were not informed in advance that some aspects of the configurations might repeat. They were simply given instructions on what the target and distractors looked like and which button to press in response to a given target. Observers were told to respond as quickly as possible without sacrificing accuracy. The experiment took approximately 50 min to complete.

*Model Results*

Before running the behavioral experiment, we first used the model to predict the results of this experiment. The model was not modified except for the parameters that are specific to the methods of this experiment. The model used the same setup as the behavioral experiment: 20 training blocks containing only the quadrant-predictive trials, followed by 10 test blocks containing both quadrant-predictive and nonpredictive trials, with an 8 × 8 screen matrix containing three items per quadrant. Data presented below are a result of averaging together 12 runs of the network, each emulating 1 observer.

The training block data are shown in Figure 7, and the test block data are shown in Figure 8. Overall, the model results predicted significant contextual cuing. The model was significantly faster at finding the target in the quadrant-predictive configurations than in the nonpredictive configurations,  $t(11) = 3.13, p < .01$ .

*Behavioral Results*

Two observers were excluded from analysis because of a large number of errors, both of which fell outside the 95% confidence interval around the mean error rate. For the other 12 observers, error rates were extremely low, around 1.5% for both nonpredictive and quadrant-predictive conditions. No difference in error rate was observed between conditions,  $t(11) = 0.464, p = .65$ .

There was a main effect of block, as shown by a one-way analysis of variance (ANOVA),  $F(19, 220) = 3.21, p = .0001$  (see Figure 9). Observers showed a significant downward trend throughout the training period as a result of practice at the task.

The mean response time was calculated separately for the two conditions in each block of the test period. In all of the experiments presented, response times of less than 200 ms or more than 6 s were discarded as outliers. This resulted in the exclusion of less than 1% of the data in each experiment. The mean was then taken over the entire test period for each of the conditions.

The test session results are shown in Figure 10. The 95-ms difference between the conditions was significant,  $t(11) = 2.73, p = .0195$ . The benefit for the quadrant-predictive condition shows that observers were able to localize targets more efficiently in this condition compared with the nonpredictive trials, demonstrating contextual cuing (Chun & Jiang, 1998).

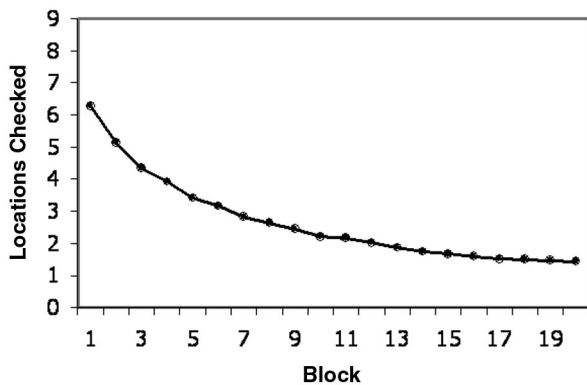


Figure 7. Model training data from Experiment 1.

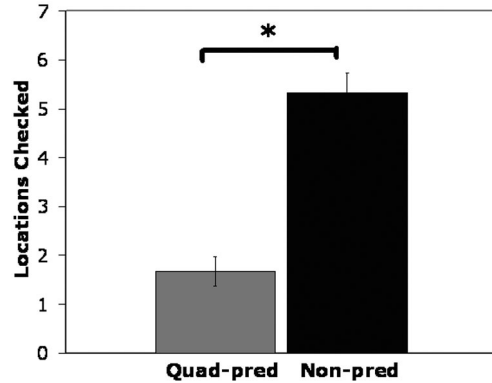


Figure 8. Model data from the test blocks of Experiment 1. The model predicts significant contextual cuing (after a significant learning period of 20 blocks). Only the blocks from the test period are shown. Error bars represent the standard error of the mean. Quad-pred = quadrant predictive; Non-pred = nonpredictive. \*  $p < .05$ .

Observers were at chance in the explicit target location guessing task, designed to test if observers were consciously aware of how contexts predicted target location. In the quadrant-predictive condition, observers guessed the correct quadrant 26% of the time, whereas in the nonpredictive condition, they guessed the correct one 20% of the time,  $t(11) = 1.26, p = .23$ . Neither condition differed significantly from chance guessing levels of 25% ( $t_s < 1.45, p_s > .12$ ). This indicates that contextual cuing was implicit.

The results of the model and the behavioral data indicate that a contextual cuing effect can be obtained from remarkably minimal predictive information embedded amidst substantial noise. In this experiment, the locations of only 2 local distractors out of 11 were predictive of the target location, and observers nonetheless demonstrated a significant benefit compared with a baseline in which no distractors were predictive of the target location. This strongly supports the idea of spatial constraints on learning, as implemented in the model. In sum, the present results suggest that a large amount of the learning seen in globally predictive contexts may be local to the target area. The degree to which this is true is further addressed in Experiment 2.

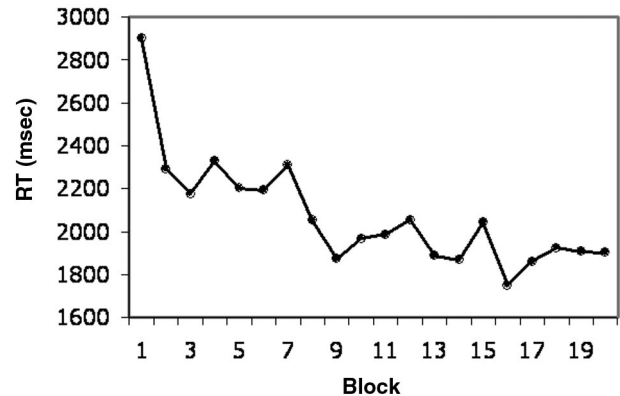


Figure 9. Training data for Experiment 1. The only condition is quadrant predictive during the training period. RT = response time.

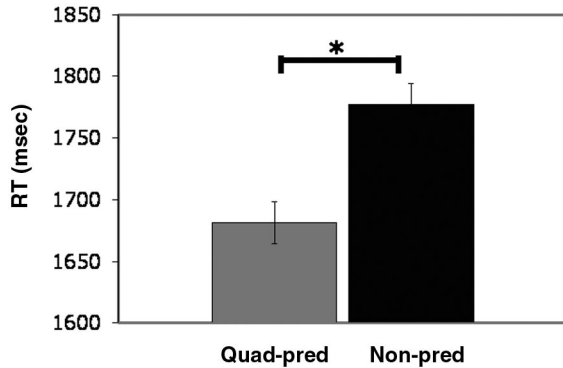


Figure 10. The results of Experiment 1. The benefit for the quadrant-predictive condition is a result of contextual cuing. Error bars represent within-observer standard error. Quad-pred = quadrant predictive; Non-pred = nonpredictive; RT = response time. \*  $p < .05$ .

## Experiment 2

Experiment 2 tested how much of the learning typically seen in contextual cuing is local to the target area. In order to do this, we trained observers on both quadrant-predictive and global-predictive contexts and looked at the magnitude of contextual cuing obtained in both cases. If observers are not learning a significant amount of information outside the target quadrant, we would expect the difference in cuing between the quadrant-predictive and global-predictive conditions to be negligible. However, if observers are learning a significant amount of information outside the target quadrant, we would expect to see a benefit for global-predictive contexts compared with quadrant-predictive contexts.

### Method

All methods were identical to Experiment 1 except where noted.

In this experiment, there were three conditions: nonpredictive, quadrant predictive, and global predictive. Training blocks consisted of 8 quadrant-predictive and 8 global-predictive configurations, and test blocks consisted of 8 quadrant-predictive, 8 global-predictive, and 16 nonpredictive configurations. To rule out target location probability effects, we used the same target locations in the nonpredictive trials as in the global- and quadrant-predictive trials.

### Model Results

Before running the behavioral experiment, we used the model to predict the results. The model was identical to that used in Experiment 1, with the addition of global-predictive trials. As in the behavioral experiment, there were 20 training blocks containing only the quadrant-predictive and global-predictive trials, followed by 10 test blocks containing quadrant-predictive, global-predictive, and nonpredictive trials. Training data were similar to that for Experiment 1, decreasing exponentially in both global- and quadrant-predictive conditions, and will not be shown for this and subsequent experiments. Data presented below are a result of averaging together 24 runs of the network and taking an average across all test blocks.

The results shown in Figure 11 indicate that the model learns equally well when only the quadrant around the target is predictive compared with when the entire screen is predictive. It does not show a significant difference between these two conditions,  $t(23) = 0.19, p > .10$ , but shows a large benefit for both compared with the nonpredictive contexts ( $ts > 2.80, ps < .01$ ). In other words, the model is doing most or all of its learning in the target quadrant. To some degree, this reflects the fact that targets are more likely to be adjacent to distractors in the same quadrant than to distractors in different, abutting quadrants. In the general case, the model may benefit slightly more from global predictive displays, simply as a result of the local context sometimes including items outside the target quadrant.

### Behavioral Results

Twenty-four observers participated in this experiment. Error rates remained extremely low at around 2% for all conditions. No difference in error rate was observed between conditions during the training period,  $t(23) = 0.29, p = .77$ , or the test period ( $F < 1$ ).

The mean response time was calculated separately for the two conditions for each block of the training period (see Figure 12). A two-way ANOVA, with block and condition as variables, revealed a main effect of block,  $F(19, 460) = 12.13, p < .001$ , but no main effect of condition,  $F(1, 460) = 1.262, p = .26$ , and no interaction ( $F < 1$ ). This suggests that participants improved over time with practice but did so equally in both global- and quadrant-predictive displays. The apparent numerical difference at Block 1 reflects random error, and it was not significant ( $p = .29$ ).

The mean response time was calculated separately for the three conditions in each block of the test period. As shown in Figure 13, there was significant contextual cuing relative to the nonpredictive condition in both the quadrant-predictive ( $M = 99$  ms),  $t(23) = 2.73, p = .012$ , and global-predictive ( $M = 169$  ms),  $t(23) = 4.37, p = .001$  conditions. The difference in cuing between the two ( $M = 71$  ms) was not significant,  $t(23) = 1.20, p = .25$ . The apparent numerical difference was the result of two outlier observers—excluding these 2 observers from the analysis reduced this

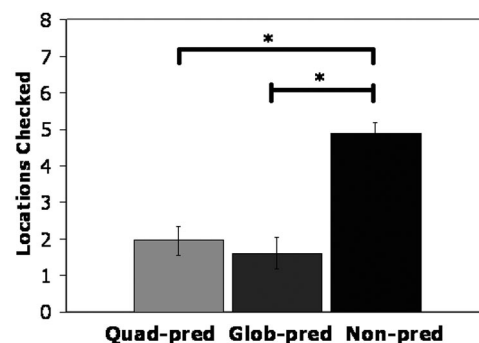


Figure 11. Model data from Experiment 2. The model predicts cuing in both quadrant-predictive and global-predictive conditions, with no significant difference between the two. Only the blocks from the test period are shown. Error bars represent the standard error of the mean. Quad-pred = quadrant predictive; Glob-pred = global predictive; Non-pred = nonpredictive. \*  $p < .05$ .



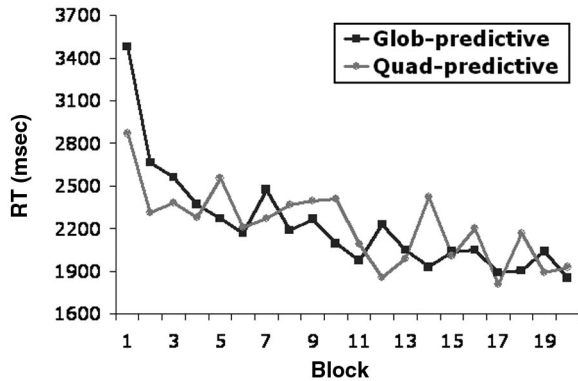


Figure 12. Training data from Experiment 2. The lack of difference between global-predictive and quadrant-predictive trials indicates that both were learned equally well. Glob = global; Quad = quadrant; RT = response time.

difference between global-predictive and quadrant-predictive conditions to under 20 ms, whereas contextual cuing for both remained significant. In sum, observers showed a statistically equivalent benefit for configurations in which only the target quadrant was repeated and for configurations that repeated entirely.

In the explicit target location guessing task, observers were at chance in all conditions at judging where the target should appear. In the quadrant-predictive condition, observers guessed the correct quadrant 22% of the time; in the global-predictive condition, they guessed correctly 27% of the time; and in the nonpredictive condition, they guessed correctly 23% of the time. None of the conditions differed significantly from chance guessing ( $t_s < 1.20$ ,  $p_s > .25$ ). This indicates that contextual cuing was implicit.

The results of this experiment indicate that most or all of the learning in the global-predictive condition was actually restricted to the target quadrant. Observers demonstrated statistically equivalent contextual cuing in both the quadrant-predictive and global-predictive conditions, indicating that they received no significant benefit from the presence of predictive information outside the target quadrant. Although it remains possible that some observers may benefit from more broadly distributed global information, the majority of observers appeared to rely on a local strategy, as predicted by the spatial constraints included in the model.

The findings so far show that contextual cuing is even more locally constrained than initially reported (Olson & Chun, 2002), and Experiment 2 replicates the lack of further advantage from global displays. These results raise two issues that were not tested in prior work, and we will describe and test each of these problems in Experiments 3 and 4.

One limitation of our conclusions is that it is possible that the training protocol of this experiment may have encouraged local processing. The target quadrant was always predictive, whereas global information was noisy on half of the displays, so an optimal (but implicit) strategy would have been to focus only on the local quadrant to minimize distraction and noise. If so, then observers should learn global information beyond the target quadrant when all training displays are global predictive. This possibility is addressed in Experiment 3.

### Experiment 3

In Experiment 2, the target quadrant was always predictive, whereas global information was noisy on half of the displays. This may have caused observers to learn only the local information in an attempt to minimize noise and optimize learning. Therefore, in this experiment, all training was done using global-predictive contexts. During the test period, half of these displays became predictive only in the target quadrant, and half remained global predictive. If observers were learning more than the target quadrant during the training period, we would expect a benefit for displays that remained global predictive compared with displays that became quadrant predictive. However, if observers were only learning the local information during the test period, we would expect an equivalent benefit for both during the test period.

### Method

All methods were identical to Experiment 1 except where noted.

Training blocks consisted of 16 global-predictive configurations, and test blocks consisted of 8 quadrant-predictive, 8 global-predictive, and 16 new-quadrant-predictive configurations. The test conditions were derived as follows. Sixteen global-predictive trials shown during the training phase were divided into two separate test conditions: half (8) remained global predictive, and half (8) became predictive only in the target quadrant. To control for learning of the quadrant-predictive configurations during the test phase, we used a baseline that was also quadrant predictive. That is, 16 new-quadrant-predictive configurations were generated using the same target locations as the global-predictive trials shown during the training phase. This controlled target location probabilities.

### Model Results

Once again, we used the model to predict the results before running the behavioral experiment. The model was identical from Experiments 1 and 2 and simulated the same setup as the current behavioral experiment: 20 training blocks containing only the global-predictive trials, followed by 10 test blocks containing

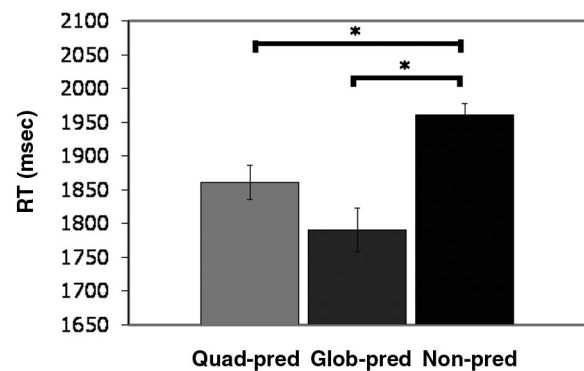


Figure 13. Results of Experiment 2. There was a significant benefit for both quadrant-predictive (Quad-pred) and global-predictive (Glob-pred) trials compared with nonpredictive trials (Non-pred) and no significant difference between them. Error bars represent within-observer standard error. RT = response time. \*  $p < .05$ .

quadrant-predictive, global-predictive, and new-quadrant-predictive trials. Data presented below are a result of averaging together 24 runs of the network and taking an average across all test blocks.

The model results shown in Figure 14 indicate that even when the entire configurations were predictive during training, only the local configurations were learned. This is indicated by the lack of a difference between the quadrant-predictive and global-predictive conditions in the test phase,  $t(23) = 0.28, p > .10$ . Both of these conditions showed a large benefit compared with the new-quadrant-predictive contexts ( $ts > 2.91, ps < .01$ ) as a result of learning during the training phase. These results confirm that the model is doing most or all of its learning in the target quadrant in this experiment.

**Behavioral Results**

Twenty-five observers participated in this experiment. One of these observers was excluded from analysis because of a high error rate, which fell outside the 95% confidence interval around the mean error rate. Error rates for the other observers were extremely low at around 2.5% for all conditions. No difference in error rate was observed between conditions during the test period ( $F < 1$ ).

The mean response time was calculated for each block of the training period. There was a main effect of block, as shown by a one-way ANOVA,  $F(19, 460) = 11.64, p < .00001$ . As in the prior two experiments, observers showed a significant downward trend throughout the training period (not shown here).

The mean response time was calculated separately for the three conditions in each block of the test period, and Figure 15 shows the results. Observers showed significantly stronger contextual cuing in both the quadrant-predictive ( $M = 73$  ms),  $t(23) = 3.16, p = .004$ , and global-predictive ( $M = 99$  ms),  $t(23) = 2.86, p = .009$ , conditions, compared with the new-quadrant-predictive baseline. The difference in cuing between the two conditions ( $M = 26$  ms) was not significant,  $t(23) = .36, p = .72$ . Thus, the benefit in both conditions was statistically equivalent. Those configurations that remained globally predictive throughout the entire experiment

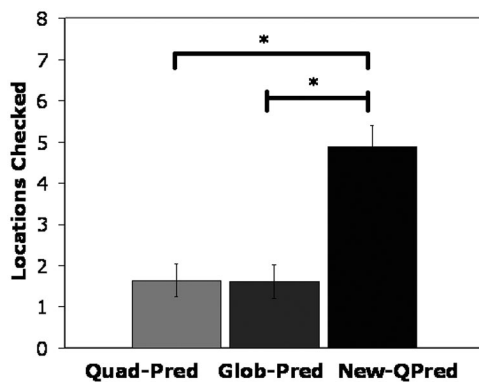


Figure 14. Model data from Experiment 3. The model predicts cuing in both quadrant-predictive (Quad-Pred) and global-predictive (Glob-Pred) conditions compared with the baseline new-quadrant-predictive condition (New-QPred), with no significant difference between the two. Error bars represent the standard error of the mean. \*  $p < .05$ .

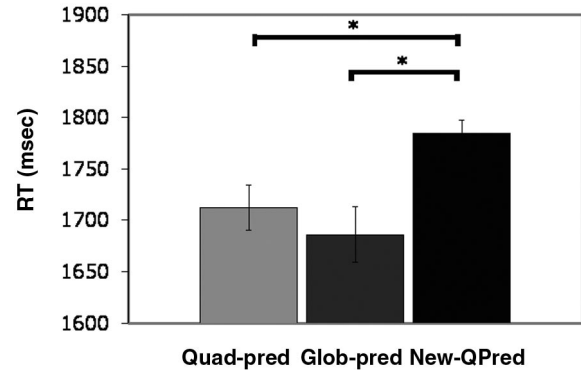


Figure 15. Results of Experiment 3. There was a significant benefit for both quadrant-predictive (Quad-pred) and global-predictive (Glob-pred) compared with the new-quadrant-predictive (New-QPred) baseline, and no significant difference between them. Error bars represent within-observer standard error. RT = response time. \*  $p < .05$ .

provided no advantage over those that were globally predictive during training but only locally predictive during the test blocks.

Observers were at chance level in all conditions at judging where the target should appear. In the quadrant-predictive condition, observers guessed the correct quadrant 23% of the time; in the global-predictive condition, they guessed correctly 24% of the time; and in the new-quadrant-predictive condition, they guessed correctly 24% of the time. None of the conditions differed significantly from chance guessing ( $ts < 1.04, ps > .30$ ). This indicates that contextual cuing was implicit.

The results strongly suggest that most or all of the learning was centered locally around the target even when the entire configuration was predictive of the target location. The lack of difference between the global-predictive and quadrant-predictive conditions suggests that the results of Experiment 2 were not a result of a change in strategy when presented with some displays that were not predictive outside the target quadrant. Thus, the learning in the contextual cuing paradigm appears to be spatially constrained to the region around the target, justifying the spatial constraints encoded in the model. Given the local nature of contextual learning, an interesting question is whether this learning is position specific or position invariant—in other words, whether it would transfer across different target locations. This question was tested in our final experiment.

**Experiment 4**

Experiment 4 was very similar to Experiment 1. The only change was that the target quadrant was allowed to move to a different quadrant of the display from repetition to repetition (see Figure 16). That is, the target quadrant configuration remained fixed, but its location on the screen and relative to the global configuration was variable from block to block.

Our motivation for this experiment was to see whether the localized learning we saw in Experiments 1, 2, and 3 was position invariant or position specific (e.g., whether it depends on the global location of the local configuration). On the surface, we would expect much of our visual learning to be position invariant; after all, people regularly learn to recognize objects even when

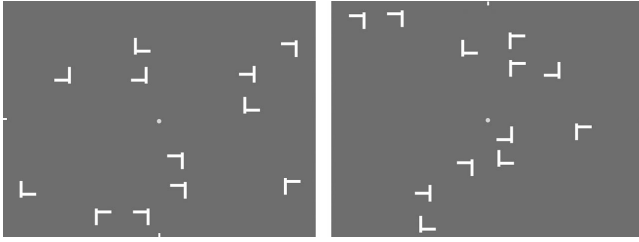


Figure 16. Example displays: In one block, the quadrant-predictive trial would look like the display on the left; in the next, it might look like the one on the right. Notice that the configuration of distractors is the same within the target quadrant (top right quadrant in the first display; lower left quadrant in the second); it has simply moved to a different quadrant of the screen.

their position in space changes. However, despite individuals' ability to recognize an object even when its projection on the retina varies significantly, several studies have shown that perceptual learning is not necessarily invariant to translation in the visual field (Dill & Fahle, 1997; Fahle, 1994; Fahle, Edelman, & Poggio, 1995; Fiorentini & Berardi, 1981; Karni & Sagi, 1991; Nazir & O'Regan, 1990; Ramachandran, 1976; Shiu & Pashler, 1992). In fact, response times and error rates tend to increase linearly with the degree of translation applied to a stimulus, which is taken as evidence that some type of shifting process is responsible for lining up visual input with stored memory (Foster & Kahn, 1985). Dill and Fahle (1997), for example, trained observers to recognize particular novel stimuli (bilaterally symmetric  $6 \times 6$  matrices of dots) and discriminate them from distractors. They then translated the stimuli to a different visual field location and found that the learning did not transfer if the training was restricted to a particular location in the visual field. This suggests that one's normal tolerance to translation may occur from exposure to familiar objects at different locations in the visual field as opposed to translation of learning to different locations in the visual field. This would indicate that we should not expect learning in this type of contextual cuing experiment to transfer between quadrants.

On the other hand, transfer of learning to translated displays in contextual cuing has been demonstrated previously, albeit with much simpler translations. In Jiang and Wagner's (2004, Experiment 2) study, contextual cuing was preserved when the entire configuration of items was translated or rescaled after learning, involving changes in the locations of the items relative to the observers but not relative to each other. However, as just noted, a significant difference is that all of Jiang and Wagner's manipulations were relative to the observer rather than relative to other items on the screen. In other words, they did not change the configuration of items on the screen, whereas we did so here. The successful transfer of learning in the Jiang and Wagner task does point to the possibility of successful learning in this experiment, however.

### Method

All methods were identical to Experiment 1 except where noted.

Experiment 4 was very similar to Experiment 1. The only change was that the target quadrant was allowed to move to a different quadrant of the display from repetition to repetition. That

is, the target quadrant configuration remained fixed, but its location on the screen was variable from block to block.

To rule out target location probability effects, we moved the target quadrant for nonpredictive trials during the test phase as well so that a given target location was occupied with equal probability for both the quadrant-predictive and nonpredictive conditions.

### Model Results

Before running the behavioral experiment, we used the model to predict the results. The model was identical to that used in all of the previous experiments, and it simulated the same setup as the behavioral experiment: 20 training blocks containing only the quadrant-predictive trials, followed by 10 test blocks containing both quadrant-predictive and nonpredictive trials. The target quadrant was moved between blocks in both conditions. Data presented below are a result of averaging together 12 runs of the network and then taking an average across all of the test blocks.

As shown in Figure 17, when participants were presented with this task, our model showed no significant contextual cuing,  $t(11) = 0.22$ ,  $p > .10$ . The predictive information about the target's location in this experiment did not benefit the model in its search for the target. This is because the model links each stimulus to its absolute location within the configuration and therefore does not generalize its learning to other quadrants.

It is important to note here that the results of Jiang and Wagner (2004, Experiment 2) discussed above do not pose a problem for our model of contextual cuing. Because our model presupposes that learning in contextual cuing depends only on the particular configuration of items and not on their relation to the observer, the model would see the translated and rescaled displays of Jiang and Wagner as identical to the original displays. It would therefore correctly predict cuing in the Jiang and Wagner task, although it predicts no cuing for the present task in which the configuration of items changes.

### Behavioral Results

Twelve observers participated in this experiment. Error rates remained extremely low, around 2% for both nonpredictive and

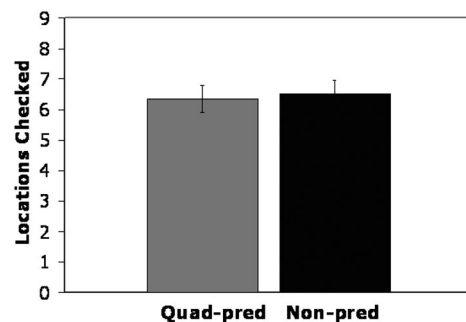


Figure 17. Model data from Experiment 4. The model predicts no cuing (even after a significant learning period of 20 blocks). Error bars represent the standard error of the mean. Quad-pred = quadrant predictive; Non-pred = nonpredictive.

quadrant-predictive conditions. There was no difference in error rate between conditions,  $t(11) = 0.146$ ,  $p = .89$ .

There was a main effect of block during the training period,  $F(19, 220) = 5.019$ ,  $p < .0001$ . Observers showed significant improvement throughout the training period (not shown).

As shown in Figure 18, the quadrant-predictive condition and the nonpredictive condition were not significantly different,  $t(11) = 0.90$ ,  $p = .38$ . In other words, observers showed no benefit for repeated configurations and were not able to localize the target more effectively when the predictive quadrant was variable in location.

Observers were at chance in both conditions at judging where the target should appear. In the quadrant-predictive condition, observers guessed the correct quadrant 26% of the time, whereas in the nonpredictive condition, they guessed the correct one 22% of the time,  $t(11) = 1.54$ ,  $p = .16$ . Neither condition differed significantly from chance guessing (both  $t_s < 0.94$ , both  $p_s > .37$ ).

The results of this experiment indicate that the quadrant-predictive contexts were not learned when the target quadrant moved location across repetitions. Observers did not benefit from the predictive information contained in such configurations, such that no contextual cuing was observed. This lack of contextual cuing indicates that the local contextual learning is somewhat tied to absolute spatial location; it is position specific rather than position invariant. This is in accordance with the model's implementation of local spatial constraints, which requires the location of items to stay the same in order for learning to occur. It is also indicative of the fact that not only is the local context important in contextual cuing but also its place relative to the global context. In other words, the locations that are learned in contextual cuing are not relative to the target location but are instead absolute to the configuration.

With more training, we would expect both human observers and the connectionist model to be able to learn in this experiment, by learning each quadrant separately. However, with the same amount of training as in Experiment 1, neither the model nor human observers benefit from the predictive context in this experiment.

The conclusions above may be restricted to spatial contextual cuing situations. When spatial configuration is predictive of target location, distractor identities are not encoded (Chun & Jiang,

1998). However, distractor identity can cue target identity if they are consistently paired through training, even when spatial configuration and target location are randomized (Chun & Jiang, 1999). Yet, such object contextual cuing may show spatially local effects as well for distractor and target identity pairings. For example, Hoffman and Seibald (2005) demonstrated that targets appearing with the same flanking distractors were detected better than when the targets appeared with rare flankers. Of note, this learning occurred independent of target position for items that could appear anywhere in a global circular array that was constant for all trials. Thus, object contextual cuing may exhibit local effects as well, but such local effects may transfer to different locations.

## Modeling Results

The model was accurate in predicting the results of the four previous experiments, and we further tested it on a broad range of published contextual cuing studies. Correctly modeling these prior experiments is critical, because the model was explicitly designed to model only two experiments—the original experiment by Chun and Jiang (1998, Experiment 1) and the first experiment to demonstrate the necessity of spatial constraints, performed by Olson and Chun (2002, Experiment 1). Remarkably, the same model is able to account for an even broader range of past findings in the contextual cuing literature, as will be detailed below. In addition, our modeling yields new insights into the operations of contextual cuing, such as why it never produces perfect attentional guidance in human observers.

### Configurations or Individual Locations

Jiang and Wagner (2004, Experiment 1) demonstrated that recombined displays benefit observers as much as completely preserved old displays. They trained observers on two sets of old displays that shared a common set of target locations (such that each target location was associated with two unique sets of distractors) and then tested them in three conditions. In their new condition, the target location was all that was preserved from training, and the display was otherwise completely novel. In their old condition, the display was one of the two displays that had been associated with the target location during training. In the recombined condition, the display was made up of a combination of distractor locations from the two displays that had been associated with the target location during training. We ran the model on this same experiment, using 11 items per display and a  $12 \times 8$  screen matrix with 20 blocks of training and 3 blocks of test, just as in the Jiang and Wagner (2004, Experiment 1) experiment; the results are presented in Figure 19.

Jiang and Wagner (2004, Experiment 1) found that observers were cued as much by the recombined displays as they were by the old displays at test. Without any modification, our model also predicted this result. This is because the model functions by associating any given input node with each output node that has served as a target when it was active. Decisions about target location are then made on the basis of how many of the active input nodes activate a particular output node location—the greater the number of active input nodes associated with a particular location, the more likely the model is to select it. This measure does not differ between the old displays and the recombined

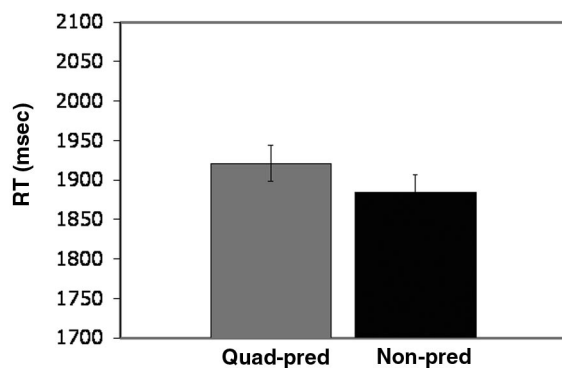


Figure 18. The results of Experiment 4. The two conditions were not significantly different, showing a lack of learning of the quadrant-predictive (Quad-pred) condition. Error bars represent within-observer standard error. Non-pred = nonpredictive; RT = response time.

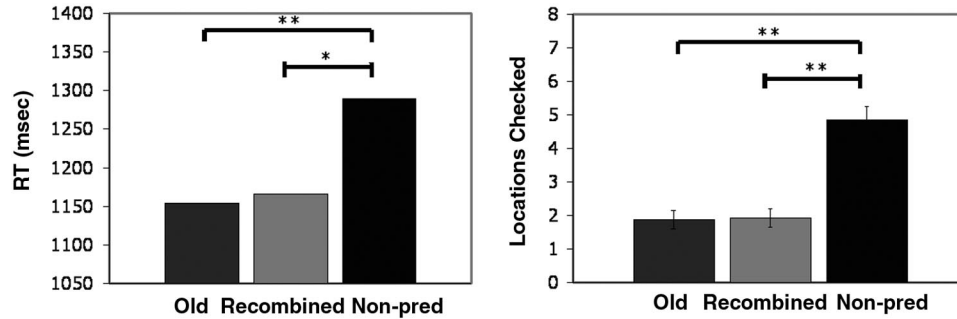


Figure 19. Data from the Jiang and Wagner (2004) experiment with recombined displays. Human data are on the left, and model data are on the right. Model error bars represent the standard error of the mean. Non-pred = nonpredictive; RT = response time. \*  $p < .05$ . \*\*  $p < .01$ . Data in the left panel are from “What is Learned in Spatial Contextual Cuing—Configuration or Individual Locations?” by Y. Jiang and L. C. Wagner, 2004, *Perception & Psychophysics*, 66, p. 457. Copyright 2004 by Psychonomic Society. Adapted with permission.

displays in this experiment—both have the same number (11) of input nodes that activate the correct target location, and therefore the model is cued equally by both types of displays. This implies that observers in a contextual cuing experiment are not encoding configurations as patterns. Instead, they are encoding the locations of multiple individual items, just as our model is doing.

The model also correctly predicts several results from the original Chun and Jiang (1998) article that it was not explicitly designed to capture. Chun and Jiang (1998, Experiment 2) demonstrated that what is learned in contextual cuing tasks is spatial configuration information rather than specific item identities. In their experiment, observers learned the context and then halfway through the experiment the distractor identities were changed. If observers had been encoding identity information, we would expect this change in distractor identities to have impaired their learning. In fact, they showed no impairment, demonstrating that spatial contextual cuing is insensitive to identity information. The model operates only over the configuration of items (as it is given only the location of the items, not their identity, as input) and so predicts cuing regardless of the particular identities of the items, paralleling the results of this experiment.

As another attempt to rule out low-level priming effects of display repetition, Chun and Jiang (1998, Experiment 3) allowed target locations to vary within the repeated contexts. If participants learned to simply search through repeated displays more quickly, search should benefit regardless of target location. However, if the predictive, associative relation is what is important, then contextual cuing should be diminished when repeated contexts were no longer predictive of target locations. Chun and Jiang (1998) found that contextual cuing was abolished when target locations varied within repeated contexts (see also Wolfe, Klempe, & Dahlen, 2000). The model replicated this result: The mean number of locations checked was 4.9 in the case in which both the display and the target location varied and was 4.8 in the case in which the display remained the same but the target location was allowed to vary within it, a nonsignificant difference,  $t(9) = 0.24$ ,  $p > .10$ . This is because cuing in the model results from the increased activation of the formerly cued target location compared with the other locations that have items. Thus, moving the target within a repeated configuration will result in no benefit compared with generating an entirely new configuration. Contextual cuing, in

both observers and the model, is based on an association between the target location and the distractor locations.

#### Learning Across Intervening Noise

Another experiment that the model can correctly predict is Experiment 3 of Olson and Chun (2002). This is a particularly important experiment because it indicates that learning in contextual cuing cannot be strictly local. In this experiment, Olson and Chun (2002, Experiment 3) once again repeated only the half of the distractors on the opposite side of the screen as the target. However, this time they broke this long-range predictive condition into two different conditions. In one condition, no noise (extraneous distractors) ever appeared between the target location and the opposite side of the display, so that all of the distractors on the same side of the screen as the target were kept farther from the target location. In the other condition, noise always appeared between the target and the opposite side of the display. These were referred to as *long-range-space* and *long-range-noise*, respectively. They found that observers were cued by the long-range distractors in the long-range-space condition but not in the long-range-noise condition. This result could not be explained if learning is strictly local. However, our model’s account of spatial constraints is able to take advantage of the distractors that are outside the target quadrant when there are fewer items close to the target to interfere with the learning. This is a result of the model’s modulation weights, which decline exponentially with distance. When there are many items close to the target, they render any effects from items that are farther away insignificant. However, when there are fewer items near the target, the small effects of the items farther away from the target become more significant in affecting search. The results of the Olson and Chun (2002) experiment and the modeling results are presented in Figure 20.

The behavioral results from Olson and Chun (2002) and the present model indicate significant cuing in the space condition,  $t(19) = 2.65$ ,  $p < .01$ , but no significant cuing in the noise condition,  $t(19) = 0.52$ ,  $p > .10$ . In the model, this result can be thought of as the spread of attention to items farther from the target when there are fewer items near the target, something one might expect in an account of learning being constrained by attentional

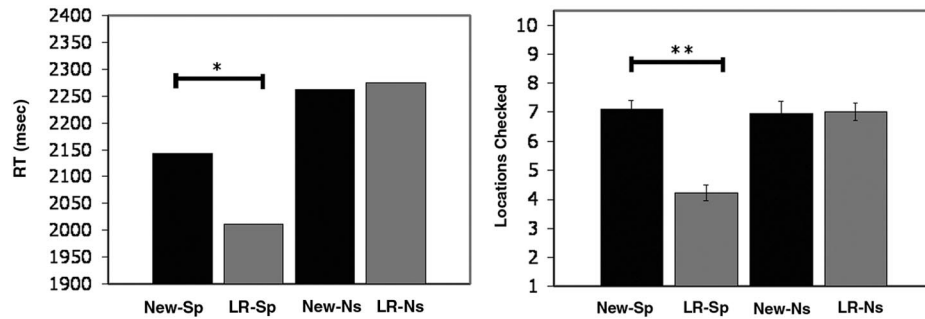


Figure 20. Data from the Olson and Chun (2002, Experiment 3) experiment. Human data are on the left, and model data are on the right. Both graphs are based on the mean of each condition in the second half of the experiment (last 10 blocks). The only comparisons made in either graph were between the new-space (New-Sp) condition and the long-range-space (LR-Sp) condition and between the new-noise (New-Ns) condition and the long-range-noise (LR-Ns) condition. Error bars on the right represent the within-subject standard error of the mean. RT = response time. \*  $p < .05$ . \*\*  $p < .01$ . Data in the left panel are from “Perceptual Constraints on Implicit Learning of Spatial Context,” by I. R. Olson and M. M. Chun, 2002, *Visual Cognition*, 9, p. 281. Copyright 2002 by Psychology Press ([www.psypress.co.uk/journals.asp](http://www.psypress.co.uk/journals.asp)). Adapted with permission.

and computational limitations. This will be considered further in the General Discussion.

#### Search Efficiency: Effects of Set Size

Likewise, the model captures the results of set size demonstrated in Chun and Jiang (1998, Experiment 4). That is, the model does not demonstrate a difference in intercept between the predictive and nonpredictive displays in search. The only difference between the two conditions in the model comes from a benefit of searching fewer and fewer items in each display as it is learned. That is, the model originally searches approximately half of the items in the display in both the nonpredictive and predictive conditions. With learning over time, the model begins to constrain its search to probable target locations that have been associated with predictive distractors.

It is worth noting that no contextual cuing experiment has demonstrated search slopes that approach zero in the predictive condition—an effect one would expect if learning contexts cued attention straight to the correct target location after suitable training. Both our model and the behavioral experiments presented above (Experiment 1–3) suggest that the lack of perfect cuing could be explained by the spatial constraints on learning. Because observers must get to a local region of the display that they have learned before they can benefit from the predictive context, it is unlikely that observers could ever reliably demonstrate search slopes near zero.

To fully explicate this idea that spatial constraints prevent contextual cuing from reaching highly efficient search slopes ( $< 10$  ms/item), we ran a probabilistic simulation of contextual cuing. This simulation was normative, in that it assumed that observers had perfectly learned the information within a local spatial window and that observers always looked at only locations that contained items. With these constraints, we estimated how the size of the spatial window influenced how quickly simulated observers would be guided to the target. By varying the set size of the displays as well, we were able to estimate and compare the search slopes in the predictive and nonpredictive conditions. Intuitively, the larger the

spatial window, the more efficient guidance should be, leading to shallower search slopes. The narrow window revealed in our experiments above suggests that search slopes will asymptote well above perfect guidance.

We generated 128 random contexts, half of which contained 8 items and half of which contained 16. Then for each context, we sampled (5,000 times) how many locations were examined when search was random—without any guidance or learning; this obviously results in about half of the display being searched on average. We then used the same contexts and sampled how long it took to find the target with perfect guidance according to the spatial constraints of our connectionist network. Whenever the model focused on a particular location, it took the probability of being immediately cued to the target location to be  $e^{-d}$  (where  $d$  was the distance between the current location and the target location), the same spatial constraints used in the connectionist model.

In addition to using the spatial constraints used by our connectionist model, we also tested several other spatial parameters to get an idea of what the actual spatial extent of learning in contextual cuing might be, based on the slopes empirically obtained from contextual cuing experiments. We tested the model with spatial windows of various widths as pictured in Figure 21.

We used this simple simulation to estimate search slopes for an observer that used no information other than learned context to search the display. The results are presented in Figure 22, along with behavioral search slopes from the last epoch (6) of the Chun and Jiang (1998, Experiment 4) contextual cuing experiment. Chun and Jiang’s (1998) set size manipulation yielded search slopes of 35 ms/item for nonpredictive (new) trials and 27 ms/item for predictive (old) trials, yielding a predictive–nonpredictive slope ratio of 76%. These ratios are used to facilitate comparison between the human behavioral response time data and the simulation results below.

The best spatial constraints for emulating contextual cuing seem to be the constraints our connectionist network is using,  $e^{-d}$ . Thus, the spatial constraints used by the model seem to have the correct shape (exponential decay) as required to emulate Olson and Chun

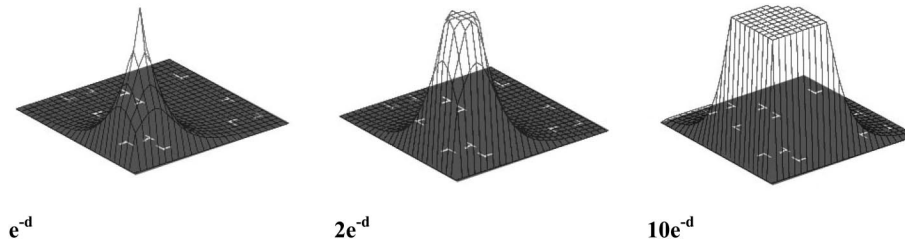


Figure 21. Graphs of the three sets of spatial constraints used, shown for the target location in these pictures (so  $d$  is always the distance between the current location and the target location). The maximum height in all of them is 1, which represents perfect cuing from a given location to the target.

(2002, Experiment 3) and the correct spatial extent as required to yield a similar gain in search efficiency as shown by Chun and Jiang (1998, Experiment 4).

Similarly, as the spatial extent of the constraints grows to encompass more of the display, this normative simulation of contextual cuing suggests that perfect cuing should be possible, as search slopes approach zero. For example, the largest spatial window simulated here produces a search benefit ratio of 12%, which would correspond to a highly efficient target search slope of 4 ms/item for predictive trials, assuming a 35 ms/item slope for nonpredictive trials. Thus, a major reason that the slope does not approach zero in contextual cuing is the restricted area of learning predicted by the spatial constraints. In fact, it suggests that it should be able to make empirical predictions about the contextual cuing benefit in search slope on the basis of how wide or narrow an individual's spatial window of attention is during the search task. For example, experimental manipulations that widen the spatial window of attention should increase contextual cuing, whereas manipulations that narrow the window should decrease contextual cuing. Of note, Lleras and Von Muhlenen (2004) showed that an active search strategy, which may lead to more spatially focused scrutiny of a display, eliminated contextual cuing. It would also be useful to see if individual differences in the

spatial extent of attentional processing may also explain variability in contextual cuing effects.

### Multiple Target Locations

Chun and Jiang (1998, Experiment 6) demonstrated that it is possible to obtain contextual cuing of two target locations with a single display but that this results in less cuing than normal. In their experiment, they paired a given set of distractors with two different possible target locations. On some trials with a given set of distractors, the target appeared in one location, and on others, it appeared in the second paired location. There was contextual cuing for both target locations, but it was weaker than that obtained when a given display always predicted the same target location. This result is also true in the model—associating two different target locations with a single display still allows learning, but search times are slowed because the active input nodes are indicative of two different output nodes being possible targets. Following the original experiment, we ran 20 blocks with a nonpredictive condition and a predictive condition. In the predictive condition, the target appeared at each of the two possible target locations 10 times randomly throughout the blocks. In the final block, the number of locations searched was 4.8 in the nonpredictive condition and 2.7 in the predictive condition,  $t(15) = 3.13$ ,  $p < .01$ , indicating that the model was able to associate at least two target locations with each display. However, cuing for two potential target locations resulted in a decrease in the predictive condition to only 2.7 locations checked, whereas our model of the original Chun and Jiang (1998) experiment that had only one potential target location decreased to 1.4 locations checked after the same training. This is similar to the effect found in this study, with a lower magnitude of contextual cuing with more target locations.

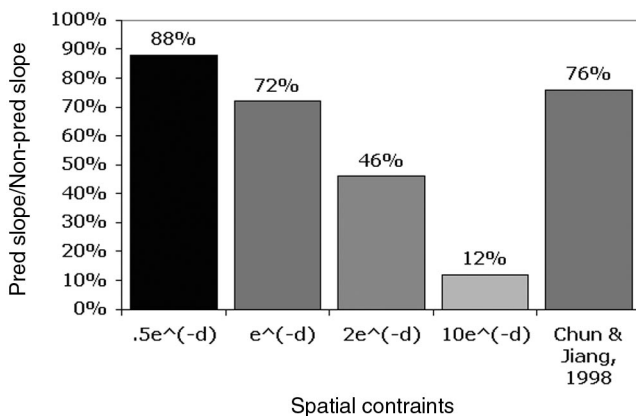


Figure 22. The difference between the nonpredictive slope and the predictive slope, expressed as a percentage. The lower the percentage, the greater the slope benefit provided by contextual cuing. Nonpredictive slopes were always 0.50 locations checked per item in the simulation. Pred = predictive;  $d$  = distance.

### Effects on Accuracy for Brief Displays

Another interesting result is that of Chun and Jiang (1998, Experiment 5). In this experiment, observers were first trained on predictive and nonpredictive displays as normal. However, the test period consisted of flashing the displays onto the screen (for 200 ms) and having observers attempt to answer about the target's direction (left or right). The important measure of contextual cuing in this task was therefore accuracy—participants were more likely to find the target on the first glance in predictive displays than in nonpredictive displays, resulting in better accuracy at the predictive trials. To represent this task in the model, we first trained the

model normally on the predictive and nonpredictive displays. We then allowed the model only one guess about the target location on each display and compared its accuracy between predictive and nonpredictive displays. As expected, the model found the target more often at the first location it checked in the predictive than the nonpredictive displays (85% of the time vs. 36% of the time),  $t(17) = 3.90$ ,  $p < .01$ , demonstrating the accuracy effect found in Chun and Jiang (1998, Experiment 5).

### General Discussion

We developed a simple model of contextual cuing using a two-layer neural network. The assumptions of the model were then tested against a novel set of experiments as well as published work on contextual cuing. Both the experiments and the fit of the model to previously performed experiments suggest that the model is an accurate portrayal of contextual cuing in several important ways. First of all, it suggests that spatial constraints on learning are needed to accurately model contextual cuing. Second, it suggests that contextual cuing is based on pairwise associations between the target location and the distractor locations that are paired with it, rather than on a gestalt perception of the global context as a whole.

#### *Spatial Constraints on Contextual Cuing*

Our model of contextual cuing used spatial constraints on learning, based on our interpretation of the results of Olson and Chun (2002, Experiment 1). We tested the specific constraints we had implemented and their predictions in Experiments 1–3. Experiment 1 demonstrated significant contextual cuing even when the predictive context was quite minimal, defined by only two neighboring distractors in the same quadrant as the target. This illustrates that observers preferentially encode the context of stimuli near the target to a degree that allows significant contextual cuing from very limited and highly localized configurations, even more localized than previously demonstrated by Olson and Chun (2002, Experiment 1).

The question we tried to answer in Experiments 2 and 3, therefore, was whether there would be any additional benefit from global context or whether observers benefit from only the local information, as the spatial constraints of our model predicted. Both Experiments 2 and 3 demonstrated that contextual cuing from local context was as strong as that observed from global context, even when observers were trained with globally predictive displays alone. This finding indicates that observers learned only the local information and is directly in line with the idea of spatial constraints that we implemented in the model.

The implementation of spatial constraints in our model of contextual cuing can be usefully described in terms of attention and the attentional spotlight, as alluded to earlier (Broadbent, 1982; Posner, 1980; Posner, Snyder, & Davidson, 1980). Contextual cuing is a form of statistical learning in which participants are encoding the statistical regularities of the display over many trials and are using it to benefit their search for the target item. Attention is necessary for statistical learning in visual search (Jiang & Chun, 2001) and in rapid serial visual presentation streams (Baker, Olson, & Behrmann, 2004; Turk-Browne, Junge, & Scholl, 2005). Specifically, observers only learned contextual and statistical associations among stimuli that were selectively attended; the asso-

ciations between unattended stimuli were not learned (Baker et al., 2004; Jiang & Chun, 2001; Turk-Browne et al., 2005). Thus, the present results suggest that attention is spatially focused around the target, encoding only a small area of information around it. Instead of encoding the entire global context throughout the visual search trial, observers may be encoding just one snapshot of the local context surrounding the target when it is detected. We also discussed in an earlier section how such local constraints may explain why contextual cuing does not produce perfect guidance.

Selective attention to the local region neighboring the target makes sense ecologically as well. It seems more important to pay attention to the local context of an item than to the larger context in most cases. For instance, recognizing a dishwasher immediately constrains the items one sees in it to be dishes or silverware, but recognizing that one is in a kitchen leaves open many more possibilities for what a group of objects in it might be. Objects may covary more frequently with items in their local context than with items in their global context, such that the local context more closely constrains what a given object might be than does its global context (cf. Hoffman & Seibald, 2005).

An even stronger incentive for local contextual learning is that computing statistical regularities is a computationally intensive process and cannot be applied to all possible relationships in the entire visual field. It makes sense, then, that spatial statistical learning should be restricted to a local region that is associated with success in the task (i.e., target detection). Likewise, these computational considerations are consistent with the results of Olson and Chun's (2002) third experiment in which there was contextual cuing from long-range context in the absence of distractors near the target: When there is no information nearby the target, learning may extend further in an attempt to grasp the context of an item. Altogether, we suggest that although computational constraints may drive most contextual learning to be spatially local around the focus of attention, the spatial extent of contextual encoding may vary depending on stimuli and task characteristics. However, we believe that a useful approximation to the type of learning seen in contextual cuing tasks is the set of fixed weights we implemented in our model, which seems capable of predicting all of the results on the spatial extent of contextual cuing to date. An interesting question for future research would be to study how spatial constraints may or may not influence contextual encoding in real-world scenes.

#### *Association and Contextual Cuing*

Another important point about the model is that its implementation via a neural network means that it learns by associating particular distractor locations with the target location. These locations are absolute locations within the configuration; they are not relative to the target location. This is a strong assumption of the model, and it predicts that learning in contextual cuing depends on an implicit association between the target's location in the configuration and the locations of the distractors within the configuration, with no relationships between the distractors' locations being learned. This form of associative learning explains why contextual cuing tolerates combined, rescaled, or displaced displays in Jiang and Wagner's (2004) experiments but not displays in which local context moves around within a global configuration (Experiment 4).



In Jiang and Wagner (2004, Experiment 1), two different displays that have both been trained to facilitate the same target location were recombined with some distractor locations from each to form a new display. This new display results in entirely preserved contextual cuing, a result our model predicts because it ties objects to their specific locations within the configuration. Decisions about target location in the model are made on the basis of how many of the active input nodes activate a particular output node location, something that does not differ between the old displays and the recombined displays in this experiment. In other words, because the specific distractor locations still code for the correct target location, cuing is preserved.

In Experiment 1 presented here, visual search was facilitated by holding the configuration of the target quadrant constant between blocks. Experiment 4 revealed that no learning occurred when the invariant quadrant was allowed to shift locations from block to block. This implies that not only is the local context important for such contextual cuing but also its place relative to the global context. This is consistent with the perceptual learning literature, which indicates that learning may not be translation invariant (Dill & Fahle, 1997; Ramachandran, 1976). If a shifting process is responsible for lining up visual input with stored memory (Foster & Kahn, 1985), then in order for target quadrant learning to transfer in the present experiments, it would be necessary to look at all possible translations for each quadrant and compare the matches with all previously seen configurations. It is easy to see why this would become computationally infeasible for the large target quadrant translations tested here. Likewise, our model predicts this result because it requires the configuration of items to be the same in order to preserve contextual cuing. Changing the location of items relative to one another will always eliminate cuing in the model, because it encodes location based on absolute location in the configuration, rather than as relative to the target location.

On the other hand, transfer of learning to translated displays like in Jiang and Wagner (2004, Experiment 2), in which the entire configuration of items was translated or rescaled after learning, is not problematic for our model. This is because what is changed is the location of the items relative to the observers but not the locations of the items relative to each other. This results in a preserved configuration of items and the preserved location of each item within the configuration, which is all the model uses to determine contextual cuing.

Therefore, the model implies that learning in contextual cuing is based on pairwise statistical associations between the distractor locations near the target and the target location. This has important implications for how such information is stored and how statistical regularities are parsed out of the incoming information. It also relates contextual cuing more strongly to the statistical learning literature, where automatic parsing of pairwise associations has been demonstrated previously, in both children and adults, and in vision and audition (Fiser & Aslin, 2001, 2002a, 2002b; Saffran, Aslin, & Newport, 1996).

## References

- Backhaus, A., Heinke, D., & Humphreys, G. W. (2005). Contextual learning in the selective attention for identification model (CL-SAIM): Modeling contextual cuing in visual search tasks. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition—Workshops—Volume 03 (June 20–26, 2005)*. Washington, DC: Institute of Electrical and Electronics Engineers Computer Society.
- Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science, 15*, 460–466.
- Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception, 25*, 343–352.
- Biederman, I. (1972, July 7). Perceiving real-world scenes. *Science, 177*, 77–80.
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143–177.
- Boyce, S. J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 556–566.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.
- Broadbent, D. E. (1982). Task combination and selective intake of information. *Acta Psychologica, 50*, 253–290.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547.
- Cave, K. R. (1999). The FeatureGate model of visual selection. *Psychological Research, 62*, 182–194.
- Chun, M. M. (2000). Contextual cuing of visual attention. *Trends in Cognitive Science, 4*, 170–177.
- 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. (1999). Top-down attentional guidance based on implicit learning of covariation. *Psychological Science, 10*, 360–365.
- Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 224–234.
- Chun, M. M., & Wolfe, J. M. (2001). Visual attention. In B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 272–310). Oxford, United Kingdom: Blackwell.
- Dill, M., & Fahle, M. (1997). The role of visual field position in pattern-discrimination learning. *Proceedings of the Royal Society of London, Series B: Biological Sciences, 264*, 1031–1036.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.
- Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology, 3*, 674–698.
- Fahle, M. (1994). Human pattern recognition: Parallel processing and perceptual learning. *Perception, 23*, 411–427.
- Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research, 35*, 3003–3013.
- Fiorentini, A., & Berardi, N. (1981). Learning of grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision Research, 21*, 1149–1158.
- Fiser, J. Z., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499–504.
- Fiser, J. Z., & Aslin, R. N. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 458–467.
- Fiser, J. Z., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, USA, 99*, 15822–15826.
- Foster, D. H., & Kahn, J. I. (1985). Internal representations and operations in the visual comparison of transformed patterns—Effects of pattern

- point-inversion, positional symmetry and separation. *Biological Cybernetics*, *51*, 305–312.
- Gibson, E. J. (1963). Perceptual learning. *Annual Review of Psychology*, *14*, 29–56.
- Gibson, E. J. (1991). *An odyssey in learning and perception*. Cambridge, MA: MIT Press.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Haykin, S. (1998). *Neural networks: A comprehensive foundation*. Englewood Cliffs, NJ: Prentice-Hall.
- Heinke, G. W., & Humphreys, G. W. (2003). Attention, spatial representation, and visual neglect: Simulating emergent attention and spatial memory in the selective attention for identification model (SAIM). *Psychological Review*, *110*, 29–87.
- Hoffmann, J., & Sebald, A. (2005). Local contextual cuing in visual search. *Experimental Psychology*, *52*, 31–38.
- Hollingworth, A. (2006). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*, *14*, 781–807.
- Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, *127*, 398–415.
- Humphreys, G. W., & Muller, H. J. (1993). SEArch vis Recursive Rejection (SERR): A connectionist model of visual search. *Cognitive Psychology*, *25*, 43–110.
- 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., & Chun, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *54(A)*, 1105–1124.
- Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cueing—Configuration or individual locations? *Perception & Psychophysics*, *66*, 454–463.
- Kanwisher, N., & Wojculik, E. (2000). Visual attention: Insights from brain imaging. *Nature Reviews Neuroscience*, *1*, 91–100.
- Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences, USA*, *88*, 4966–4970.
- Lleras, A., & Von Muhlenen, A. (2004). Spatial context and top-down strategies in visual search. *Spatial Vision*, *17*, 465–482.
- Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, *103*, 603–649.
- Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics*, *2*, 547–552.
- Miller, J. (1988). Components of the location probability effect in visual search tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 453–471.
- Nazir, T. A., & O'Regan, J. K. (1990). Some results on translation invariance in the human visual system. *Spatial Vision*, *5*, 81–100.
- Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, *9*, 273–302.
- Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory and Cognition*, *3*, 519–526.
- Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- Peterson, M. A., & Kramer, A. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics*, *63*, 1239–1249.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*, 160–174.
- Ramachandran, V. S. (1976, July 29). Learning-like phenomena instereopsis. *Nature*, *262*, 382–384.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Shiu, L.-P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics*, *52*, 582–588.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552–564.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.
- Wolfe, J. M., Klempe, N., & Dahlen, K. (2000). Post-attentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 693–716.
- Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223–256). London: University College London Press.
- Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 601–621.

Received May 10, 2005

Revision received December 18, 2006

Accepted December 29, 2006 ■