

# Investigating Cue Competition in Contextual Cuing of Visual Search

T. Beesley and David R. Shanks  
University College London

A fundamental principle of learning is that predictive cues or signals compete with each other to gain control over behavior. Associative and propositional reasoning theories of learning provide radically different accounts of cue competition. Propositional accounts predict that under conditions that do not afford or warrant the use of higher order reasoning processes, cue competition should not be observed. We tested this prediction in 2 contextual cuing experiments, using a visual search task in which patterns of distractor elements predict the location of a target object. Blocking designs were used in which 2 sets of predictive distractors were trained in compound, with 1 set trained independently. There was no evidence of cue competition in either experiment. In fact, in Experiment 2, we found evidence for augmentation of learning. The findings are contrasted with the predictions of an error-driven associative model of contextual cuing (Brady & Chun, 2007).

*Keywords:* contextual cuing, visual search, cue competition, blocking, augmentation

Understanding the fundamental processes and properties of learning has been a central goal of experimental psychology for over a century, yet there remains a deep division within the field at the broadest theoretical level. Associative accounts posit that learning the relationship between events involves the strengthening of links between the mental representations of stimuli, through processes such as reinforcement and prediction error. Associationism has a long and successful history as an explanatory framework for learning phenomena across a wide range of species (see Bouton, 2007) and has led to a number of influential formal models of learning. A profoundly different account describes learning as a process of higher order cognition, arising through causal reasoning and the formation of beliefs (see Mitchell, De Houwer, & Lovibond, 2009).

One aspect of learning that must be accounted for by any theory is how stimuli compete for the control of behavior: Why is it that some cues are learned about at the expense of others? Kamin (1969) demonstrated that when two stimuli, A and B, are conditioned in compound (AB+, with + indicating the outcome event), learning is impaired relative to when these stimuli are conditioned apart (B+). In addition, after a cue is initially trained as a signal for reinforcement (A+), learning about a novel cue is impaired when conditioned in compound with this pretrained cue (AB+; poor learning about B). Kamin termed these cue-competition phenomena *overshadowing* and *blocking*, respectively, and there have been many subsequent demonstrations of these effects in a variety of conditioning studies across a range of nonhuman animals (see Macphail, 1982) as well as in human contingency learning (HCL)

tasks (e.g., Dickinson, Shanks, & Evenden, 1984; Le Pelley, Beesley, & Suret, 2007).

Following Kamin's (1969) findings, blocking and overshadowing became key explananda for models of associative learning. The famous model developed by Rescorla and Wagner (1972; hereinafter the Rescorla–Wagner model) provides an explanation in terms of limitations in processing of the unconditioned stimulus (the reinforcing outcome). Changes in associative strength ( $\Delta V_A$ ) in the Rescorla–Wagner model are determined by the discrepancy between the actual outcome occurring and the expectation of the outcome based on the current associative strength of all stimuli ( $\lambda - \Sigma V$ ):

$$\Delta V_A = \alpha \times \beta(\lambda - \Sigma V), \quad (1)$$

where  $\alpha$  and  $\beta$  are learning rate parameters. Increases in the associative strength of the pretrained stimulus,  $V_A$ , will lead to a decrease in the prediction error on trials in which it predicts the outcome, ( $\lambda - V_A$ )  $\rightarrow$  0. When the novel stimulus B is presented during compound trials, the combined associative strength of all stimuli,  $V_A + V_B$ , will lead to minimal prediction error, ( $\lambda - V_A + V_B$ )  $\approx$  0. Thus, as the discrepancy between the expected and actual outcomes will be small, little will be learnt about cue B.

The success of the Rescorla–Wagner model in explaining a wide range of learning phenomena has positioned it as a standard formulation of modern associative learning theory. Although the model is not without its shortcomings (see Miller, Barnet, & Grahame, 1995), the principles of learning through error correction have received support both experimentally and physiologically. For instance, activation in the dorsolateral prefrontal cortex has been shown to be sensitive to changes in associative strength following unexpected outcomes (Fletcher et al., 2001), whereas at the neuronal level, dopamine neuron responses are consistent with differences in prediction error during learning tasks (Schultz & Dickinson, 2000). Thus, prediction error is a fundamental property of the neural processes involved in learning and a strong case can be made for its role in cue competition (Waelti, Dickinson, & Schultz, 2001).

Propositional accounts, in contrast, suppose that participants make a comparison between the events occurring on different trials during training and that causal judgments arise through inferential reasoning

---

This article was published Online First September 5, 2011.

T. Beesley and David R. Shanks, Division of Psychology and Language Sciences, University College London, London, United Kingdom.

This work was supported by ESRC Grant RES-062-23-1778.

We thank Luis Jiménez for his helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to T. Beesley, Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London WC1H 0AP, United Kingdom. E-mail: t.beesley@ucl.ac.uk

processes (e.g., Mitchell, De Houwer, & Lovibond, 2009). For example, in the case of blocking, participants may note that both the probability of occurrence and the nature of the outcome on AB+ trials were the same as those they experienced on the pretraining A+ trials. Therefore, participants may infer that B is not an independent cause of the outcome, because B appears to have no effect on the properties of the outcome or its probability of occurring.

A number of studies have provided support for such a propositional account of cue competition (for reviews, see De Houwer, Beckers, & Vandorpe, 2005; Shanks, 2010). For example, De Houwer, Beckers, and Glautier (2002) used a blocking procedure in which outcomes could either occur with maximal intensity (20/20) or with submaximal intensity (10/20). Blocking effects were stronger when submaximal intensity outcomes were used, suggesting that participants were able to infer that cue B had the potential to change the magnitude of the outcome on AB+ trials (increase it to 20/20), but it did not and was therefore unlikely to be a cause of the outcome.

Of course it may well be the case that both associative and propositional reasoning accounts are correct but that their respective mechanisms may be more or less likely to be engaged by certain task conditions. Traditionally, cue competition has been examined in HCL tasks, which provide corrective feedback to allow participants to learn through a process of trial and error. During a final test phase, participants are asked to reflect on their knowledge of the task structure in order to provide causal judgments. These conditions are likely to promote the use of controlled reasoning processes during both learning and retrieval. It follows, therefore, that clearer evidence for the role of associative processes in human cue competition would come from an observation of such effects under conditions that restrict controlled reasoning processes.

These conditions are met in implicit learning tasks. In such tasks, learning is incidental, in the sense that participants are given a cover task that does not explicitly highlight the contingencies present in the task material and in the sense that measurements of learning are taken during ongoing performance rather than during offline reflective judgments. The experiments presented in this article provide an examination of cue competition in one such implicit learning task: the contextual cuing task developed by Chun and Jiang (1998, 1999, 2003). Thus, the question these experiments seek to answer can be simply stated: Is cue competition observed in incidental learning conditions, where inferential processes are unlikely to be engaged? If so, the involvement of error-driven associative learning would be more firmly established.

The contextual cuing task is ostensibly a visual search task: Participants are required to search for a target stimulus (e.g., a *T* oriented 90° to the left or right) within a set of distractor stimuli (e.g., rotated *Ls* and *Fs*). Participants are required to use left and right response keys to make a speeded response to the orientation of the *T*. Critically, some configurations of distractor and target stimuli are repeated throughout the experiment, such that it is possible for participants to use the repeating pattern of distractors to cue the target location. On other trials the distractor stimuli are randomly arranged, such that they do not predict the position of the target object. The studies of Chun and Jiang (1998, 2003) and numerous subsequent studies found that target detection was faster on repeating patterns than on novel patterns, suggesting that participants can learn to associate the repeating distractor configurations with the target location.

Unlike in HCL tasks, participants are able to perform this task with perfect accuracy without engaging in learning: They simply have to locate the target and respond to it. Thus, it is assumed that learning the configurations of repeating distractor patterns in the contextual cuing task reflects incidental learning. Note that the claim here is not that learning proceeds implicitly (unconsciously or without attention) in the contextual cuing task. Although such claims have been made repeatedly (e.g., Chun & Jiang, 1998, 1999, 2003; Goujon, Didierjean, & Marmèche, 2009; Hoffmann & Sebald, 2005; Rausei, Makovski, & Jiang, 2007; van Asselen & Castelo-Branco, 2009), recent tests of awareness using sensitive recognition and generation measures found evidence for explicit knowledge in contextual cuing (Smyth & Shanks, 2008) and evidence concerning the role of attention is similarly mixed (Jiang & Leung, 2005; Rowland & Shanks, 2006). Moreover, the evidence for associative learning processes is not directly refuted by evidence of conscious knowledge, because the issue of awareness is orthogonal to the issue of the representational level of learning operations (associative or inferential accounts; see Shanks, 2010). To summarize, the observation of cue competition in the current tasks would point toward the operation of error-driven learning processes and would be difficult to reconcile with an account of cue competition based purely on the operation of controlled reasoning processes.

### A Formal Model of Contextual Cuing

For each experimental design presented here, it was first necessary to check whether learning of the task contingencies would be expected to lead to cue-competition effects. To this end, we subjected our designs to the associative model of contextual cuing developed by Brady and Chun (2007), which is capable of producing a cuing effect by learning associations between the spatial locations of distractor elements and the spatial location of the target.<sup>1</sup>

The model is a two-layer neural network in which the input and output layers of units both represent the same stimulus locations within the display. Input units are activated by the presence of a stimulus in the represented location: If a stimulus (target or distractor) is present, then the activation of the input unit is 1; otherwise, it is 0. Activation on the output units reflects the model's anticipation of a target at these locations, which is achieved by passing the activation from the input units through a set of learned weights. Thus, on trials with repeating contexts, associations strengthen between the positions of distractor stimuli and the target position, whereas for nonrepeating trials these associations are weak.

A full description of the model is presented in Brady and Chun (2007), and some key features of the model are described in the Appendix. Here, we concentrate only on how the model learns to update the weights connecting input and output units. The learned weights are adjusted using a form of the delta rule:

$$\delta_o = \lambda - a_o, \quad (2)$$

<sup>1</sup> Kunar, Flusberg, Horowitz, and Wolfe (2007) have shown that contextual cuing might not be driven by an enhancement in the guidance of attention toward targets but rather by a reduction in the threshold required for responses to targets in cued locations. Whereas Brady and Chun (2007) clearly intended their model to reflect attentional guidance, the representational framework of the model is consistent with either account.

$$(\Delta LW_{oi})_t = \varphi \times \delta_o \times a_o \times a_i + \theta \times (\Delta LW_{oi})_{t-1}, \quad (3)$$

where  $\delta_o$  is the prediction error of output unit  $o$ , determined by the difference between the expected value of the output unit,  $\lambda$ , and the actual activation of output unit  $o$ ,  $a_o$ . In Equation 3,  $\varphi$  is a learning rate parameter determining the contribution of the current prediction error to the weight change,  $a_i$  is the activation of input unit  $i$ , and  $\theta$  is a momentum term that determines how much the previous weight change ( $t - 1$ ) contributes to the direction and magnitude of the current change ( $t$ ). In summary, the model learns through incremental weight changes that aim to minimize the error on the output units. Thus, for a repeating pattern, the input units representing the present distractor pattern will develop large weights with the actual target output unit.

Brady and Chun (2007) demonstrated by computational simulations how the model can provide an account of many effects from the contextual cuing literature, such as the spatial constraints on distractor–target learning (Olson & Chun, 2002) and unimpaired cuing following the recombination of separately trained contexts (Jiang & Wagner, 2004). Brady and Chun stressed that the model is presented as an account of contextual cuing effects only and that it is not intended to account for other important aspects of visual processing, such as the effect of target–distractor similarity (e.g., Duncan & Humphreys, 1989).

The similarity of the delta rule (Equation 2) to the Rescorla–Wagner rule (Equation 1; see Sutton & Barto, 1981) encourages the hypothesis that cues within a repeating pattern will compete for associative strength in the model: It is sufficient to have only a few highly predictive elements within a pattern (each with a strong association to the target) to produce little or no error on the target output unit. Therefore, in cases where there is zero error for a given output unit ( $\delta_o = 0$ ), no adjustments to the weights relating to this output location will occur. Consider now the behavior of the network if we were to introduce more predictive elements to a pattern that contains a set of established predictive elements. For these stimuli to acquire an association with the target location, they need some degree of prediction error in the model ( $\delta_o > 0$ ). Thus, in a blocking procedure, there is little capacity for the model to learn about additional predictive elements.

## Design and Simulation of Experiment 1

We set out to formally test whether blocking of associative learning occurs in Brady and Chun’s (2007) model. The design of the blocking procedure is detailed in Table 1. The uppercase letters in the design refer to sets of eight distractor stimuli. The lowercase letters refer to whether that particular set of distractors was predictive of the target location. If the pattern of distractors was repeated across trials, it was therefore predictive of the target location ( $p =$  predictive). Patterns of distractors that did not repeat across trials (on each trial distractors were randomly arranged) were not predictive of target location ( $r =$  random).

A schematic of this design is presented in Figure 1. The experiment uses a single-phase blocking design. During the training phase, six different trial types were used, of which four are represented in Figure 1. Pattern type Ap refers to patterns in which the distractors repeated and so were predictive of the target location. In a particular pattern, the arrangement of A distractors consistently predicted the location (but not the orientation) of the target  $T$ . For pattern type Cr,

Table 1  
*Design of Experiment 1*

Training	Test
Ap	ApBr
ApBp	ArBp
ArBr	ArBr
Cr	CpDr
CpDp	CrDp
CrDr	CrDr

*Note.* The experiment contained four sessions, each containing a training phase followed by a test phase. Uppercase letters A through D refer to different sets of eight distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r).

the distractors were randomly arranged on each trial. For trial types ApBp and CpDp, all distractors were predictive of target location. The difference between these two sets of compound distractor patterns, however, was that in the case of ApBp patterns, half the distractor elements were independently trained as predictive of the target location (Ap). Thus, the logic of the design is that training the Ap patterns separately will lead the elements of the Ap pattern to become associated with the target location. Following sufficient training, these established associations will result in minimal error on the output units. Importantly, this reduction in prediction error will occur not only on Ap trials but also on ApBp trials. Hence, pairing the Ap and Bp elements in the compound pattern ApBp should lead to attenuated learning about the Bp elements. Because neither element of the CpDp compound is trained independently, learning about these elements should not suffer the associative blocking that occurs for Bp elements.

The four test trials shown in Figure 1 were used to measure the cuing effect generated by each component of the distractor patterns. The logic of each test trial type is to test the cuing of one set of distractors while the other set is randomized and therefore not able to cue the target location. For example, on test trial ArBp, the A elements were randomly arranged such that any cuing of the target location (relative to a suitable control trial with no predictive cues at all) can be due only to the Bp elements. Test trials constructed in a similar way were used to measure the cuing generated by Ap, Cp, and Dp elements.

## Pattern Creation

Figure 2 shows example displays of actual patterns used in Experiments 1 and 2. In our experimental tasks, participants were given screens containing letters  $F$  and  $L$  or  $R$  and  $B$  (see later). However, we continue to use the letters A through D to refer to the different trial types used in the designs. For trial types Ap, ApBp, and CpDp, four individual patterns were created, each with a different configuration of distractor elements. For these four different configurations, the target stimulus was located in a different quadrant of the grid. For example, for the first ApBp pattern, the target was positioned in the top left quadrant of the grid, whereas for the second configuration of type ApBp, the target was located in the top right quadrant of the grid, and so on.

Trial types Ap and Cr featured only one set of distractors and were created by placing two distractors in random positions in each of the four quadrants of the grid. All other trial types featured two sets of distractors and were created by randomly placing two

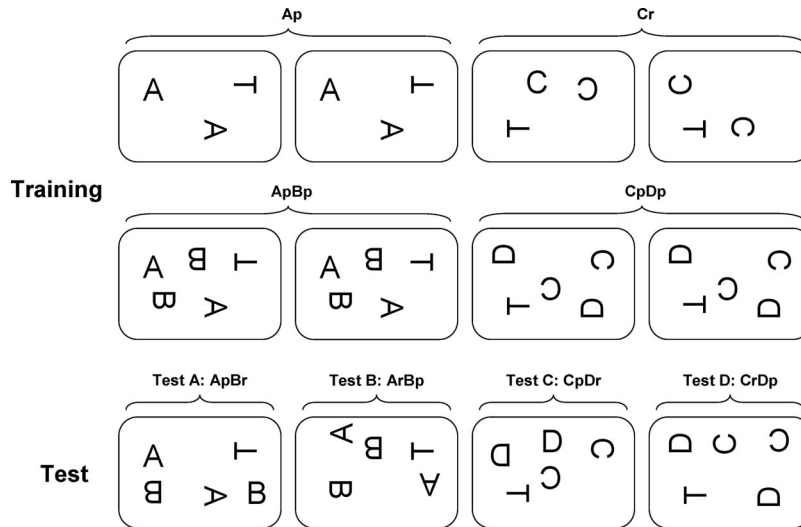


Figure 1. A schematic of the design of Experiment 1. Letters A through D refer to the design elements in Table 1. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r). T = target stimulus.

distractors of each letter set in each quadrant. ApBp patterns were created by adding the Bp elements to the preexisting Ap patterns, thus maintaining the configuration of Ap distractors across these two trial types. Patterns for ArBr and CrDr trial types were created in the same manner as for CpDp trials, but new configurations were created for each block of the experiment (these configurations did not repeat).

Twelve different target locations were used, three from each quadrant of the screen. Four target positions were used for the Ap and ApBp patterns, and four different targets were used for the Cr and CpDp patterns. Thus, the targets used for ApBp and CpDp patterns were presented an equal number of times. Four more target positions were used for the ArBr and CrDr random patterns. Target location was also constrained to a ring shaped region around the grid, such that a target was never displayed near the center of the grid or near the corners. Because the constraints on target location were applied to all trial types, biases in responding across trial types cannot be due solely to the properties of target location.

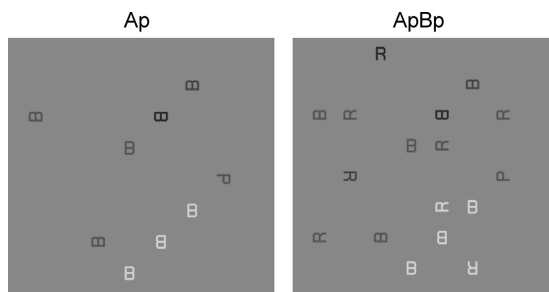


Figure 2. Example displays of trial type Ap used in Experiment 1 and ApBp used in Experiments 1 and 2 (A and B = sets of distractor stimuli; p = predictive target position). Each stimulus was presented in one of four colors. P = target stimulus.

In the test phase, two test patterns were created from each compound pattern by replacing one set of distractors with a novel random arrangement. For example, the trial type ApBr was created by replacing the B distractors with a novel configuration on each test trial. The position of the target on these trials was consistent with the original compound training pattern. Random patterns ArBr and CrDr were created in the same manner as in the training phase and used the same four target positions.

### Simulation Results

Because we were not fitting the model to any data, the chosen parameter values were similar to those used in Brady and Chun's (2007) simulations.<sup>2</sup> The model was run with 500 simulated subjects, trained with patterns generated by the procedures detailed earlier.

The model was first tested with the same amount of training as our human participants received (see Experiment 1): 16 blocks of training and six blocks of test. Figure 3 shows the results of this simulation. Figure 3A shows the level of activation for the target output unit during the 16 blocks of training, for the six trial types. Initially, activation of the target unit decreased rapidly, caused by the rapid adjustment of the initialized weight structure (small excitatory and inhibitory weights) to a vastly inhibitory weight structure.<sup>3</sup> The model then learned rapidly about predictive patterns of input. Activation of the target output unit was highest

<sup>2</sup> The learning rate ( $\varphi$ ) was set at .01, the momentum term ( $\theta$ ) was .95, the bottom-up component ( $\beta$ ) was .1, and the width of the attentional spotlight ( $\chi$ ) was set at 10. A higher learning rate parameter was used in the current simulations (.01 rather than .001), though the ordinal predictions of the model are consistent across a range of parameter values.

<sup>3</sup> Because a nonlinear activation function is used in the model (Equation 4), large negative weights must be formed between these units and active input units (i.e.,  $in_o \rightarrow -\infty$ ) for the model to produce zero activation on nontarget output units.

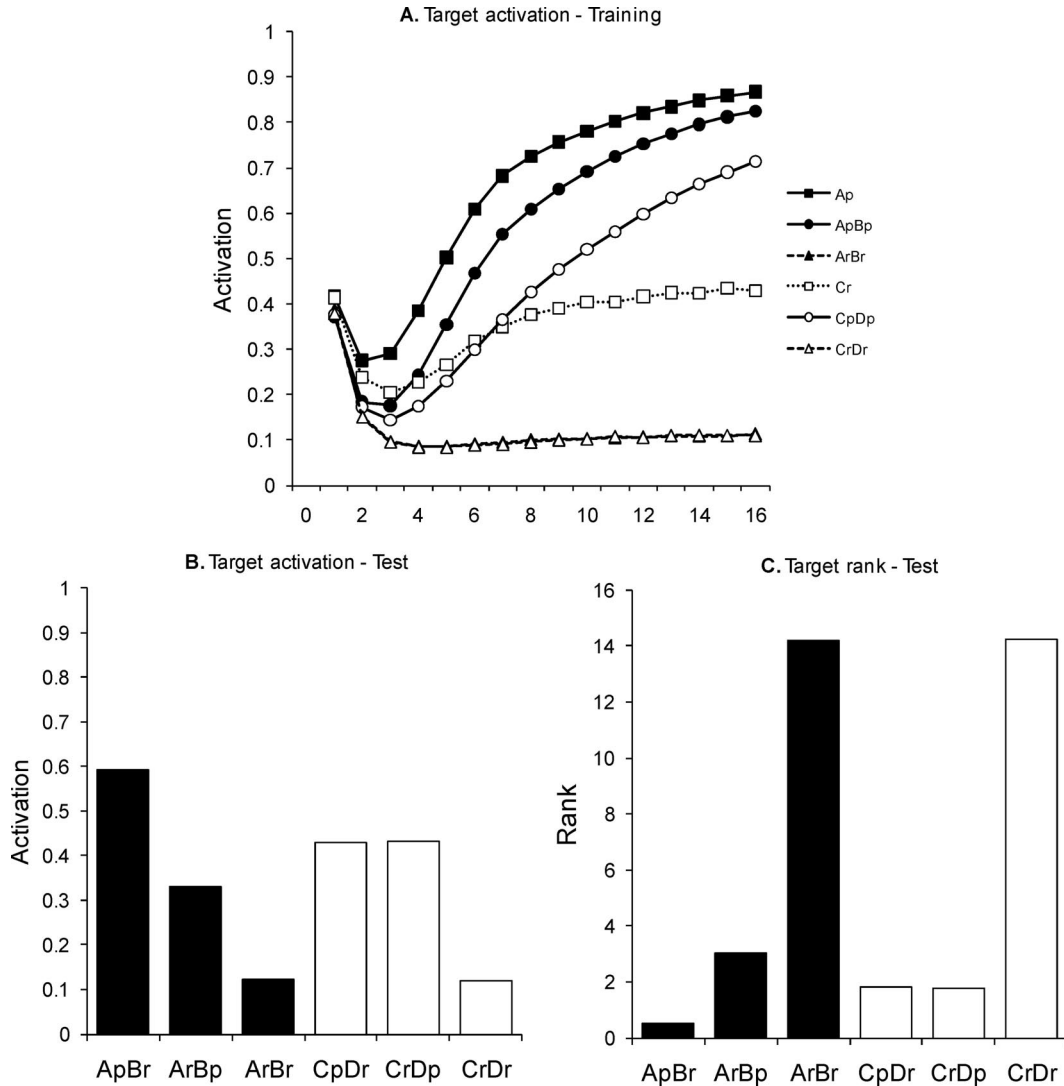


Figure 3. Simulation results of Experiment 1 from Brady and Chun’s (2007) model with 16 blocks of training. Panel A shows activation of the target unit during the training phase. Panel B shows activation of the target unit during the test phase. Panel C shows the position of the target unit among all output units ranked by activation level during the test phase. In Panels A and B, higher values reflect more efficient search performance. In Panel C lower values reflect more efficient search performance. Uppercase letters A through D refer to different sets of eight distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r).

when the network was presented with the Ap patterns, demonstrating strong learning of the distractor configurations for these patterns. Similarly high performance was seen for the ApBp patterns, presumably driven to a large extent by learning of the Ap component.<sup>4</sup> Performance on the CpDp patterns also reached a high level by the end of the training phase, whereas the network performed poorly when presented with novel, randomly arranged patterns (Cr, ArBr, and CrDr). Target activation on the Cr patterns actually increased over the course of training. This outcome reflects the consistent placement of the target on the Cr and CpDp trials: The input activation generated by the target location itself was enough to cause a moderate level of activation on the target output unit.

Figures 3B and 3C show the model’s performance during the test phase. Figure 3B shows the activation of the target output unit,

<sup>4</sup> It is likely that the slightly lower target output activation on ApBp compound patterns compared with Ap patterns reflects the greater degree of interference caused by patterns with more elements (ApBp has twice as many active elements as Ap). Many of the patterns presented to the model will contain some degree of overlap in terms of their active input elements. This overlap will more often than not cause interference in the form of predictions for an irrelevant target location. Those patterns with more active elements will contain, in total, more elements in common with other patterns; therefore, on average, interference will be greater for compound patterns.

with highest activation shown on ApBr trials in which Ap patterns were the only predictive set of distractors. Equivalent activation was generated by Cp and Dp patterns (assessed on CpDr and CrDp trials, respectively), which, given the reduced amount of training, showed lower activation levels than Ap patterns. Despite Bp patterns receiving the same amount of training as both Cp and Dp patterns, lower levels of activation were generated on the target output unit by the Bp patterns (i.e., target activation was lower on ArBp trials than on CpDr and CrDp trials). Low levels of target activation were observed on ArBr and CrDr trials, indicating that the higher levels of activation for Bp, Cp, and Dp patterns constitutes learning of these repeating patterns of context.

Figure 3C shows the test data by a different measure. Here the output units have been ranked in terms of their activation, and the y-axis shows the mean rank position of the target output unit. The pattern is inversely related to the raw target activation levels, with poorest performance (high rank values) on the all random trials and best performance (low rank values) on the test of the Ap patterns. Again, poorer performance was observed on Bp patterns compared with Cp and Dp patterns. Therefore, the patterns of results in both the raw activation data and the target rank data show that the model predicts a blocking effect for this design.

To test the performance of the model after asymptotic learning of both compounds, the model was given extended training of 100 blocks. Figure 4 shows the results of this simulation. A more pronounced blocking effect is seen compared with performance after training with 16 blocks. Thus, the model predicts that the greater the number of presentations of Ap and ApBp patterns, the greater the blocking effect will be.

In summary, the model predicts that learning about the predictive nature of a pattern of distractors is impaired if that pattern is presented in compound with distractor elements that are independently trained as predictive of the target location. In other words, the model predicts a blocking effect: slower and less efficient target search on ArBp test trials than on CpDr and CrDp test trials.

## Experiment 1

The pattern creation method used in Experiment 1 was the same as that used for the model simulations described earlier. Given that the model made the prediction that the size of the blocking effect would increase with extended training, participants were trained over the course of four sessions. Participants received a training phase followed by a test phase in each session, conducted in morning and afternoon sessions on consecutive days.

## Method

**Participants.** Sixteen participants were recruited from the psychology subject pool of University College London. Participants took part in exchange for course credit or payment of £16. Each participant was tested over four sessions, each lasting for approximately 35 min. Over the course of 2 consecutive days, one session was conducted in the morning and one in the afternoon. Morning and afternoon sessions were separated by a break of at least 2 hr.

**Materials and apparatus.** The experiment used two separate sets of stimuli, denoted by letters A and B and by letters C and D in the design (see Table 1). Trials contained either a target stimulus

letter *T* or a *P*. On trials with two sets of distractors, the target letter *T* was intermixed with *F* and *L* letters, whereas the target stimulus *P* was intermixed with *R* and *B* letters. On other trials, only one set of distractors was displayed. To control for any differences in the ease with which the target letters could be discriminated from the sets of distractors, the assignment of stimuli was counterbalanced. For half of the participants, the stimuli *F* and *L* were used for distractor sets A and B, and stimuli *R* and *B* were used for distractor sets C and D. For the remaining participants, this assignment was reversed. Furthermore, within each subset of participants, the assignment of stimuli within each set was also counterbalanced (e.g., whether *F* or *L* represented letter A), creating eight different configurations in total.

The experiment was conducted on PCs with 17-in. TFT (thin film transistor) monitors set at a resolution of 1024 × 768. Stimulus presentation and response recording was handled by software programmed with Cogent 2000 and Cogent Graphics ([www.vislab.ucl.ac.uk/cogent.php](http://www.vislab.ucl.ac.uk/cogent.php)). Stimuli were drawn within the experiment software. The letter stimuli were 9 mm wide and 13 mm high in their upright orientation. Stimuli were arranged in a square grid of 64 evenly spaced cells, which was positioned centrally on the screen and was 193 mm square. The grid was invisible to participants. The fixation cross (displayed centrally before each trial) was 11 mm square. Responses to the target stimulus (see later) were made with keys *x* and *m* on a standard PC keyboard.

The background color of the screen was gray. Stimuli were blue, red, green, or yellow. Distractor stimuli could be oriented by rotating the letter by 0°, 90°, 180°, or 270°. Target stimuli could be oriented by rotating the letter by 90° or 270°. The color and orientation of the distractors were randomly assigned for each pattern, with the constraint that there could be no more than four distractors of one color, and there had to be at least one distractor of each color, for patterns containing eight distractors. For patterns containing 16 distractors, there could be, at most, eight distractors of one color, and there had to be at least two distractors of each color. Repeating elements of patterns maintained the same color and orientation for distractor stimuli across repetitions. Target color was also randomly determined and maintained across presentations of the same pattern, but target orientation was determined randomly within each block of trials (see Figure 2 for example patterns).

**Procedure.** At the start of the first session, participants received instructions detailing the nature of the task. Example displays were presented, and participants were shown the correct response for each orientation of the two targets. Participants completed a practice block of 16 trials containing target stimuli without distractor stimuli, to allow familiarization with the appropriate responses for the different target orientations. Participants then began the main task. In Sessions 2 through 4, participants were told that the task was the same as in the previous session and received no further instructions and no practice trials. Sessions lasted approximately 35 minutes.

Each session contained 16 training blocks followed by six test blocks. Participants were not made aware of the transition between the training and test phases. Blocks contained four patterns of each trial type (see Table 1) resulting in blocks of 24 trials. Within each block, trials were presented in a random order with the constraint that consecutive trials across adjoining blocks could not present the same pattern. A rest break of 20 s was given after every four

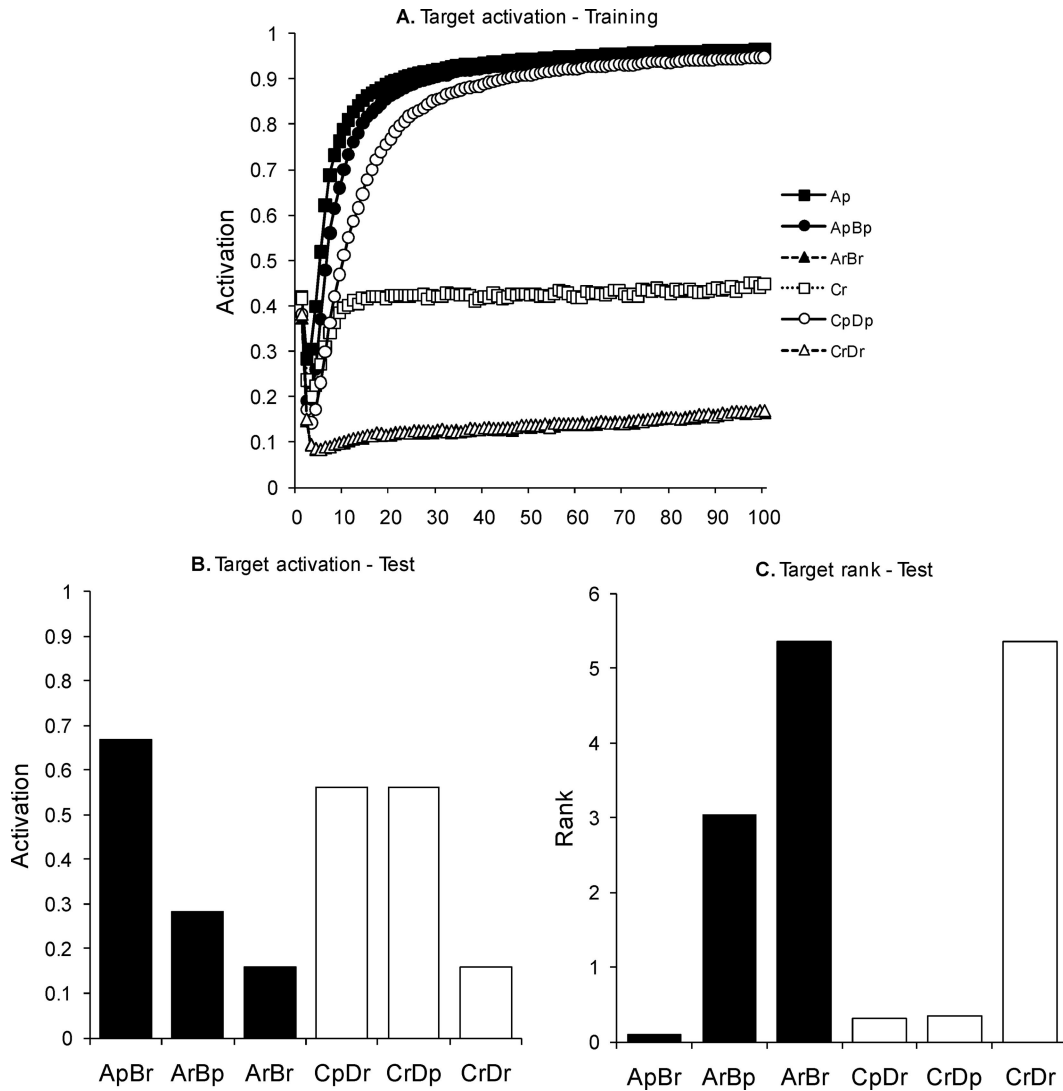


Figure 4. Simulation results of Experiment 1 from Brady and Chun’s (2007) model with 100 blocks of training. Panel A shows activation of the target unit during the training phase. Panel B shows activation of the target unit during the test phase. Panel C shows the position of the target unit among all output units ranked by activation level during the test phase. In Panels A and B, higher values reflect more efficient search performance. In Panel C, lower values reflect more efficient search performance. Uppercase letters A through D refer to different sets of eight distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r).

blocks, and trials started automatically after this period. Target orientation (left or right) was determined randomly but with an equal number of presentations of each orientation within a sub-block of eight trials.

Each trial commenced with a fixation cross presented in the center of the screen for 1,000 ms. This fixation cross was then replaced immediately by the pattern of stimuli. Reaction times (RTs) were recorded from the onset of the pattern. Following a valid response (*x* or *m*) the pattern was removed from the screen. The response–stimulus interval was 1,000 ms. If participants made an incorrect response to the target orientation, *ERROR!* appeared in the center of the screen for 2,000 ms, prior to the response–stimulus interval.

**Results**

One participant’s mean RT across all four sessions was more than three standard deviations from the mean of the sample. This participant was therefore an extreme outlier in terms of overall RT and was not included in the analysis.

The mean accuracy across the four sessions for the remaining 15 participants was 98.2%, with each session above 97.8%. Because of the clear ceiling effects, accuracy was not analyzed further.

RTs on trials on which an error occurred were not analyzed. RTs were natural log transformed (logRT) to improve the normality of the distribution. To further reduce the influence of the slowest responses, logRTs that were three standard deviations greater than

the mean were also removed (<1%), and participant data points for each trial type were calculated with median averages.

Figure 5A shows the logRTs for the six trial types across the training phases of the four sessions. The data were collapsed to four-block averages for analysis (each session actually contained 16 training blocks, with a total of 64 training blocks across the four sessions). Learning emerges early on in training for Ap, ApBp, and CpDp patterns as shown by the faster RTs for these trial types compared with their comparison trial types (Cr, ArBr, and CrDr, respectively). Because of the overall difference between RTs to patterns with eight and 16 distractors, these different trial types were subjected to separate anal-

yses. An analysis of variance (ANOVA) with factors of trial type (Ap vs. Cr) and block revealed main effects of trial type,  $F(1, 14) = 3.79, p < .05$  (one tailed), and block,  $F(15, 210) = 37.89, p < .001$ , and a significant Trial Type  $\times$  Block interaction,  $F(15, 210) = 2.78, p < .01$ . These results indicate learning across blocks for Ap compared with Cr. An ANOVA with factors of set (AB vs. CD), trial type (p vs. r), and block revealed a main effect of trial type,  $F(1, 14) = 42.21, p < .001$ , and block,  $F(15, 210) = 88.17, p < .001$ , and a significant Trial Type  $\times$  Block interaction,  $F(15, 210) = 3.62, p < .001$ . There was no main effect of set, nor did this variable interact with the factors of trial type or block, all  $F_s < 2.45, p_s > .14$ .

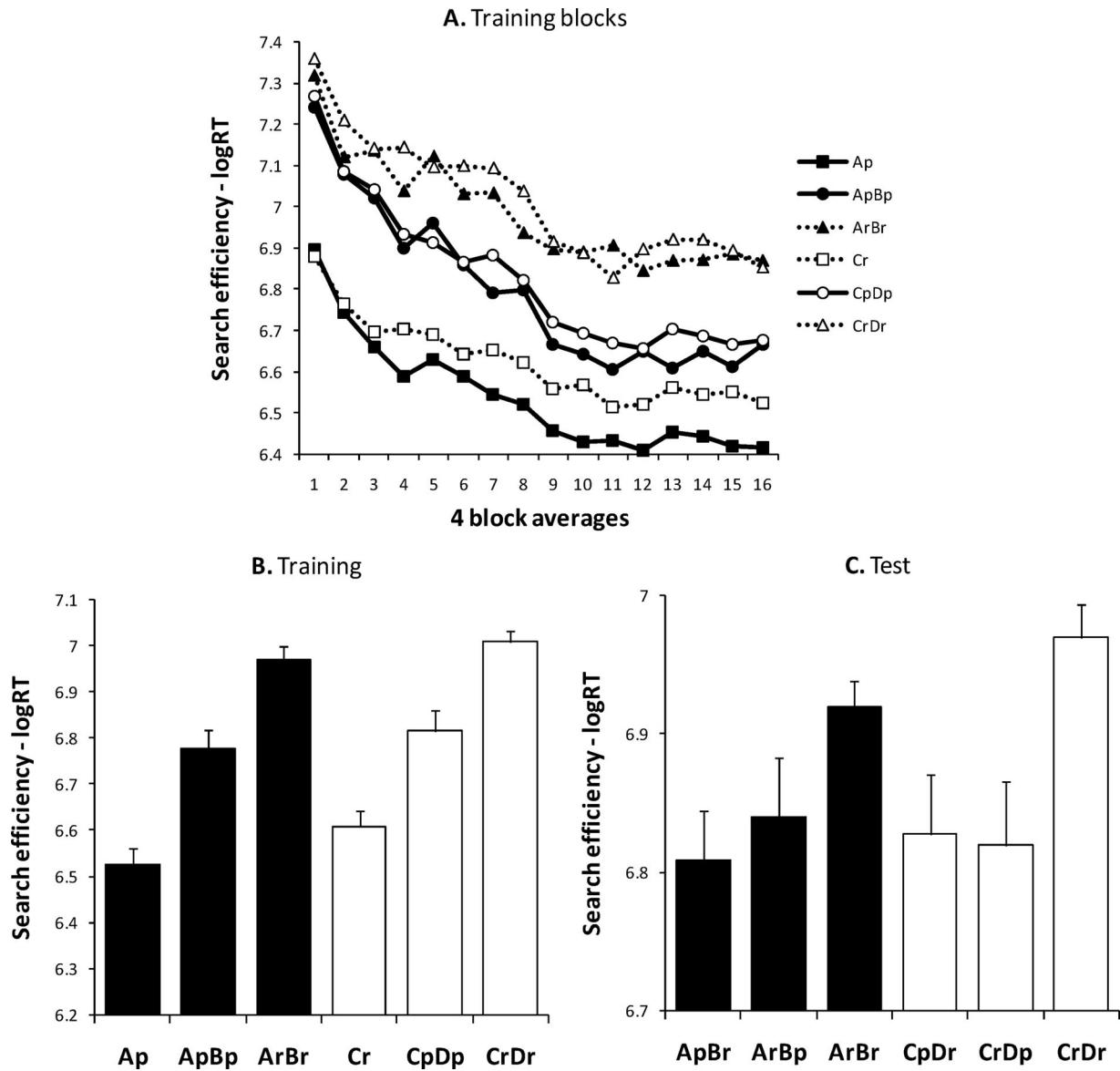


Figure 5. Target search efficiency during Experiment 1 in log transformed reaction time (logRT). Panel A shows training data collapsed into four-block averages across the four sessions. Panel B shows training data averaged across all training blocks and sessions. Panel C shows test data averaged across all test blocks and sessions. Uppercase letters A through D refer to different sets of eight distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r).



Calculating the difference in target search times for predictive trial types and their random counterparts (i.e., logRTs on Cr minus logRTs on Ap; ArBr minus ApBp; CrDr minus CpDp) provides a learning score for each set of distractors, with positive scores indicating more efficient search for predictive patterns. Analyzing each training session separately, paired-samples *t* tests revealed that there was no evidence for learning of the Ap patterns in Session 1,  $t < 1$ , or Session 2,  $t(14) = 1.71, p = .11$ , but evidence for learning did emerge in Sessions 3 and 4, both  $t(14) > 2.34, ps < .05$ . There was evidence for learning in every session for compound trial types, all  $t(14) > 2.61, ps < .05$ . Furthermore, there was no difference between the learning observed for the different compound patterns in any of the four sessions, all  $t < 1$ .

Figure 5C shows the test data averaged across all four sessions of the experiment, in which cuing was independently examined for the four predictive distractor sets (Ap, Bp, Cp, and Dp). Target search times on ApBr trials were significantly shorter than on ArBr trials,  $t(14) = 3.72, p < .01$ , supporting the findings from the training phase that learning of the Ap distractor sets had occurred. The difference between ArBp and ArBr was also significant,  $t(14) = 1.97, p < .05$  (one tailed). Target search times were also shorter on both CpDr trials,  $t(14) = 3.18, p < .01$ , and CrDp trials,  $t(14) = 2.62, p < .05$ , compared with those on CrDr trials. Directly comparing search times for predictive test patterns, pairwise comparisons between the four different test pattern types revealed no differences, all  $t(14) < 1.02, ps > .33$ .

We can assess whether learning of the Bp distractors was blocked (because of its pairing with the Ap distractors) by comparing whether the facilitation in target search time for Bp patterns (ArBr – ArBp) was smaller than that for Dp patterns (CrDr – CrDp). Although the Cp and Dp distractor sets were equally predictive of target location, Dp is the most appropriate comparison for Bp, because for each participant, the letter set used for the D elements was presented the same number of times as the set used for the B elements (whereas the letter set used for C was presented more often). A repeated measures ANOVA on these learning scores, with factors of pattern (Bp vs. Dp) and session (1–4), showed no main effects of pattern or session,  $F_s < 1$ , and no interaction effect,  $F(9, 126) = 1.06, p = .40$ . Averaging across all sessions, the learning score for Bp was .08 ( $SD = .16$ ), and the score for Dp was .15 ( $SD = .22$ ). These learning scores were not significantly different,  $t(14) = 1.09, p = .29$ . The power to detect a medium-sized effect (see Cohen, 1998) was .58 (one tailed;  $d = .5, \alpha = .05$ ).

## Discussion

An intermixed blocking schedule was used in Experiment 1. Participants received training with compounds of distractor stimuli (ApBp and CpDp), and in the case of one set of compounds, half of the distractors were also trained separately as predictive of the target location (Ap). Simulations demonstrated that this procedure produced a blocking effect—poor learning of the Bp patterns—in an associative model of contextual cuing. Empirically, learning was shown for the compound patterns during the training phase, as well as for the pretrained single pattern by the end of this phase. However, when each distractor set was tested individually, there was little evidence of cue competition.

## Experiment 2

Although it was clear that the model predicted a blocking effect when trained with the intermixed design used in Experiment 1, it is possible that certain changes to this procedure might provide a more sensitive empirical test of blocking. Specifically, the intermixed training of Ap and ApBp sets allows for an initial period in which the prediction error in the model could facilitate learning of the Bp–target associations and, hence, could reduce the size of the blocking effect. In addition, it is possible that changing from Ap patterns, which contain eight distractors, to ApBp patterns, which contain 16 distractors, might lead to a certain degree of generalization decrement. That is, the expected blocking effect may be weakened because of the inability of the ApBp patterns to activate the learned representations of the Ap patterns from training.

Experiment 2 addressed these two concerns. The experiment was run with a larger sample size and a design that we hoped would increase the size of the effect we were hoping to detect. A blocking design was used in which the Ap patterns were trained alone during the first two sessions, presented with additional random distractors (ApBr) to minimize the generalization decrement that may occur when transferring between training phases. This pretraining phase should have facilitated the development of strong associations between the Ap distractors and the target location, such that when the compound ApBp was presented subsequently, the Ap patterns would already be capable of reducing the prediction error and would act as stronger competing cues to block learning about the Bp distractor patterns.

## Design and Simulation of Experiment 2

The design of Experiment 2 is shown in Table 2. Training 1 contains compounds of distractor stimuli designed to pretrain the Ap distractors as predictive of the target location. In contrast to Experiment 1, all trials in Experiment 2 contained two sets of distractor stimuli (i.e., every display included one target and 16 distractors). The intention of this change was to make the repeating patterns of Ap distractors less obvious to participants. Therefore, by including more random materials in this stage, we hoped to make the salience of the repeating patterns of distractors comparable with the salience of repeating patterns in Experiment 1.

Table 2  
Design of Experiment 2

Training 1	Test 1	Training 2	Test 2
ApBr	ApBr (T1)	ApBp	ApBr ArBp
CrDr	ArBr (T3)	ArBr	ArBr
	CrDr (T2)	CpDp	CpDr CrDp
	CrDr (T4)	CrDr	CrDr

*Note.* The experiment contained four sessions. Sessions 1 and 2 contained phases Training 1 and Test 1. Sessions 3 and 4 contained an initial period of Training 1, followed by Training 2 and Test 2 phases. Uppercase letters A through D refer to different sets of distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r). T1 through T4 refer to different sets of target positions.

Four sets of target locations were used. The notations T1 through to T4 refer to which set of target positions was used with each distractor set. The assignment of these target positions to distractor sets is explained later.

Similar to Experiment 1, comparison of performance on ApBr and CrDr trials during Training 1 provides a measure of learning about the predictive Ap distractor sets. However, for any given participant, target detection may be easier with, say, distractor letters *F* and *L* than with distractor letters *B* and *R*. Whereas the counterbalancing of distractor stimuli will minimize the impact of these differences on performance when assessed over the whole sample, we sought a test that would provide a more accurate measure of learning at an individual participant level. Test 1 provided this measure, by comparing target search times on ApBr with those on random distractor patterns, ArBr, which were created with the same stimuli. Note here that different target positions were used for the new sets of random trials in Test 1 (ArBr with target set T3 and CrDr with target set T4). By using these untrained target sets, we were also able to measure the effect of target frequency on cuing performance; we would expect performance to be better on the trained targets, T1 and T2, than on the untrained targets, T3 and T4. Confirmatory evidence of learning about the Ap distractor set would therefore be shown by a greater difference in target search times between T1 and T3 compared with the difference between T2 and T4.

Training 2 provided training with compounds of the distractor stimuli, ApBp and CpDp. These compound trials used the same target locations as were used in Training 1 (T1 and T2). Random trials continued to use the target locations used in Test 1 (T3 and T4).

Test 2 provided the critical test of learning about predictive distractor sets A through D. The logic of each comparison is the same as that used in Experiment 1. For instance, to assess whether learning about Bp trials differed from learning about Cp and Dp trials, we can compare the difference in performance on ArBp and ArBr to the difference in performance on CpDr (or CrDp) and CrDr.

The model was trained with 48 blocks of Training 1, six blocks of Test 1, 24 blocks of Training 2, and eight blocks of Test 2. This training schedule was similar to the training structure given to participants (see later). The model was trained with the same parameters and number of simulated subjects as the simulations conducted for Experiment 1.

Figure 6A shows the activation of the target output unit over the course of Training 1. Activation was higher on ApBr trials than on CrDr trials, indicating learning of the Ap distractor patterns. As in Experiment 1, target output activation increased across training for CrDr trials, indicating some degree of learning about the fixed set of target locations for these trials. This target frequency effect was also evident in Figure 6B, which shows the results of Test 1. Here, very little output activation occurred on presentation of patterns paired with the new target positions T3 and T4, and performance on CrDr trials with T2 targets was better than on those presented with T4 targets. Hence, the model showed strong learning of predictive distractor sets and also demonstrated a target frequency effect.

Figure 6C shows target activation during Training 2. Target activation on both ApBp and CpDp trials reached a high level by the end of this training period, with higher activation seen on ApBp trials driven by pretraining of ApBr in Training 1.

Test performance is shown in the lower panels of Figure 6. Focusing on the target output activation data (Figure 6D), higher

activation was observed when the model was presented with the Ap distractor sets (ApBr). Moreover, target output activation was lower on Bp test trials than that on Cp and Dp test trials, which demonstrates a clear blocking effect in the model. As in the simulations of the Experiment 1 design, when using the activation rank measure, the corresponding pattern of results was observed (Figure 6E). That is, search performance was less efficient on Bp test trials (more locations were checked before the target was found) compared with control test trials Cp and Dp.

## Method

**Participants.** A new sample of 20 participants was recruited from the psychology subject pool at University College London. Participants took part in exchange for course credit or payment of £16. The length and organization of the four sessions were the same as in Experiment 1.

**Materials and apparatus.** The stimuli were created in a manner similar to that used in Experiment 1, but in Experiment 2, all patterns contained two sets of distractors. A ring of 16 target locations was selected, with these locations approximately an equal distance from the center of the screen. For each participant, these 16 locations were divided equally between the four different target sets (i.e., T1–T4), with the constraint that each target set had one target in each quadrant of the screen. The apparatus and stimulus creation methods were identical to those in Experiment 1 with each set of predictive distractors (e.g., Ap) comprising four repeating patterns.

**Procedure.** The method of stimulus presentation was identical to that used in Experiment 1. Sessions 1 and 2 contained 48 blocks of Training 1, followed by six blocks of Test 1. Sessions 3 and 4 began with four blocks of Training 1, which served as a reminder presentation of the Ap patterns. Participants then received 16 blocks of Training 2, followed by eight blocks of Test 2.

Each block contained four patterns from each trial type. Thus, Training 1 blocks contained eight trials, Test 1 blocks contained 16 trials, Training 2 blocks contained 16 trials, and Test 2 blocks contained 24 trials. Within each block, trials were presented in a random order with the constraint that consecutive trials across adjoining blocks could not present the same pattern. A rest break of 20 s was given after every 60th trial, and trials started automatically after this period. Target orientation (left or right) was determined randomly but with an equal number of presentations of each orientation within a subblock of eight trials.

## Results

The mean accuracy was 97.9% ( $SD = 1.8$ ), with each session above 97.4%. Data transformations and exclusions for long RTs (<1%) were conducted in the manner outlined in Experiment 1.

Figure 7A shows data from the Training 1 phase collapsed into four-block averages for the two sessions (each session contained 48 blocks of Training 1). An ANOVA with factors of trial type (ApBr and CrDr) and block revealed main effects of trial type,  $F(1, 19) = 11.88, p < .01$ , and block,  $F(23, 437) = 41.72, p < .001$ , and a significant interaction effect,  $F(23, 437) = 2.21, p < .01$ . Participants clearly learned about the repeating ApBr patterns in Training 1.

Figure 7B shows data averaged across the first two sessions and across blocks, with the data from Training 1 presented as the first

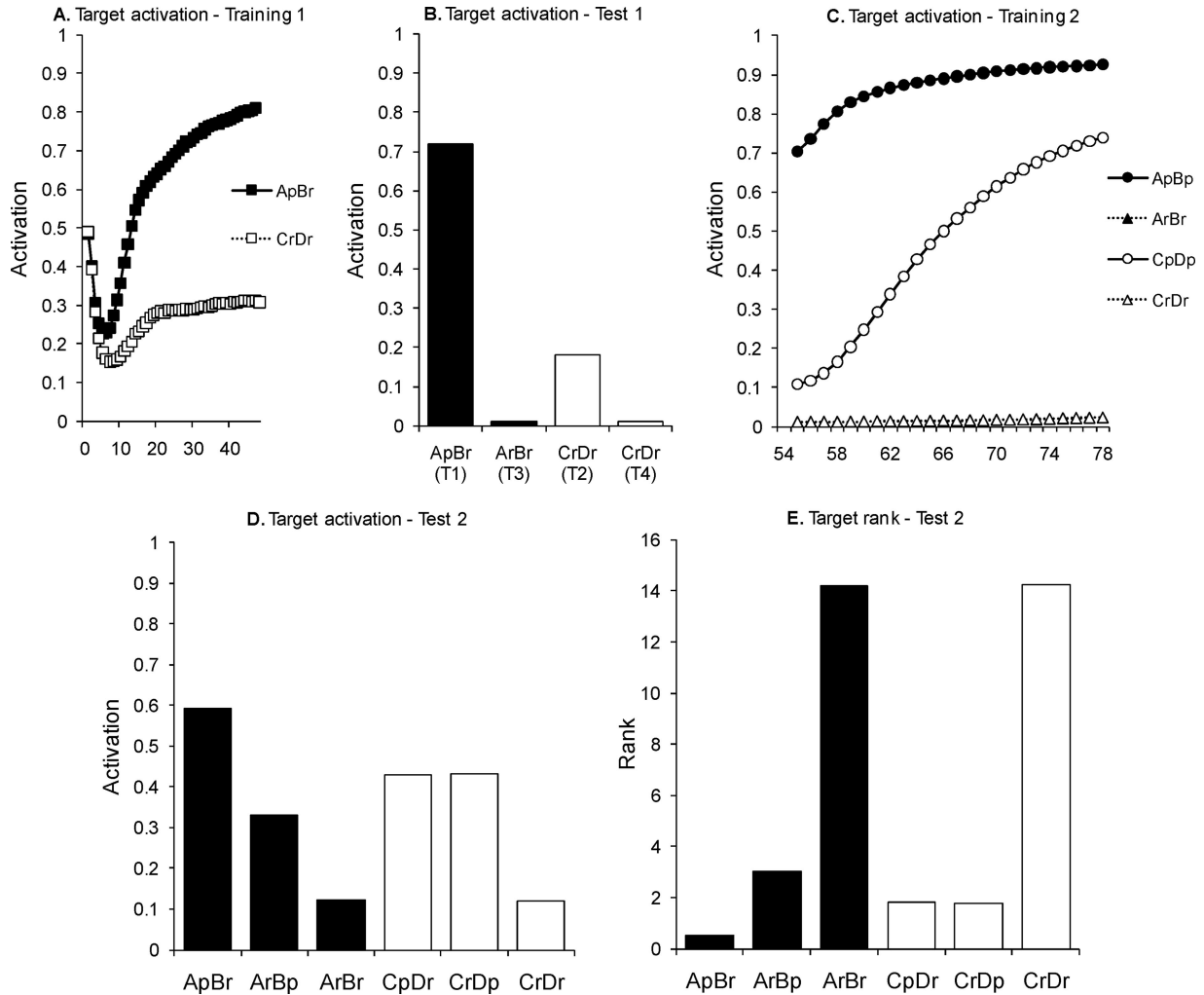


Figure 6. Simulation results of Experiment 2 from Brady and Chun’s (2007) model. Panel A shows activation of the target unit across Training 1. Panel B shows activation of the target unit during Test 1. Panel C shows activation of the target unit across Training 2. Panel D shows activation of the target unit during Test 2. Panel E shows the position of the target unit among all output units ranked by activation level, during Test 2. In Panels A through D, higher values reflect more efficient search performance. In Panel E, lower values reflect more efficient search performance. Uppercase letters A through D refer to different sets of eight distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r). T1 through T4 refer to Target Sets 1 through 4.

two bars and the data from Test 1 as the last four bars. As described earlier, an additional test of learning is provided by comparing the difference in target search times on frequent and infrequent targets for the different sets of stimuli. An ANOVA was conducted on the Test 1 data in Figure 7B (last four bars) with set (AB vs. CD) and target (T1 and T2 vs. T3 and T4) as within-subject factors. This revealed a main effect of set,  $F(1, 19) = 4.53, p < .05$ , suggesting that target search was overall more efficient with AB than CD distractors. The main effect of target was highly significant,  $F(1, 19) = 51.37, p < .001$ , which confirms that target search was faster on familiar target positions (T1 and T2) than on novel target positions (T3 and T4). More important, the interaction was significant,  $F(1, 19) = 8.80, p < .01$ , indicating that the difference in target search efficiency for familiar and novel target

sets was greater for AB patterns, which provides further evidence of the Ap distractor learning during Training 1.

Figure 8A shows data from the Training 2 phase as four-block averages for Sessions 3 and 4 (each session contained 16 blocks of Training 2). It is clear from the faster responses to ApBp compared with CpDp at the start of the Training 2 phase that there was substantial generalization between the ApBr and ApBp patterns. To evaluate whether learning of the CpDp compounds occurred, the data from CpDp and CrDr were subjected to a repeated measures ANOVA with factors of block and trial type (CpDp vs. CrDr). This revealed a main effect of block,  $F(7, 133) = 10.18, p < .001$ , and trial type,  $F(1, 19) = 20.47, p < .001$ , and, more important, a significant interaction,  $F(7, 133) = 2.09, p < .05$ . This interaction indicates that the

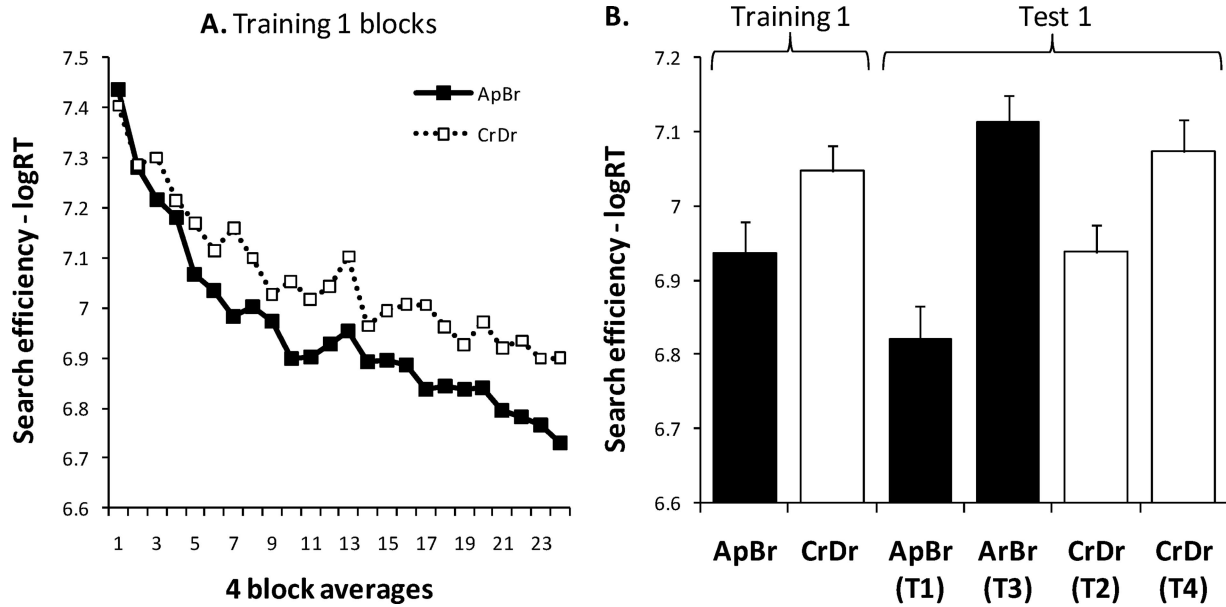


Figure 7. Target search efficiency during Sessions 1 and 2 of Experiment 2 in log transformed reaction time (logRT). Panel A shows Training 1 data collapsed into four-block averages across both sessions. Panel B shows Training 1 and Test 1 data averaged across all blocks and both sessions. Uppercase letters A through D refer to different sets of eight distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r). T1 through T4 refer to Target Sets 1 through 4.

difference in responding between the predictive and nonpredictive trial types increased across the two sessions of training and, therefore, that participants learned about the CpDp distractor patterns during this training.

Figure 8C shows the data from Test 2. It is clear that target search times on ApBr trials were shorter than those on all other trials. A paired-samples *t* test revealed a significant difference between ApBr trials and ArBp trials,  $t(19) = 2.69, p < .05$ , and between ApBr and the average of CpDr and CrDp,  $t(19) = 3.40, p < .01$ .<sup>5</sup> Crucially, however, there was no difference between target search times on ArBp trials and the average of CpDr and CrDp trials,  $t(19) = 1.45, p = .16$ , which suggests that pairing the Bp distractors with the pretrained Ap distractors did not result in competition for learning. In fact, when assessed against the baseline provided by the patterns of random distractors (ArBr and CrDr), the difference in target search times on ArBr and ArBp trials,  $M = .23 (SD = .17)$ , was greater than the difference between CrDr and the average of CpDr and CrDp,  $M = .12 (SD = .16)$ ,  $t(19) = 2.81, p < .01$ . This finding suggests that rather than competing for learning resources, pairing the B distractors with the pretrained A distractors led to augmentation of learning.

## Discussion

In Experiment 2, participants were pretrained with predictive distractor patterns, Ap, for the first two sessions, before compound training was introduced. Simulation results with Brady and Chun's (2007) model showed that this pretraining procedure produced a strong blocking effect. The data from the first two sessions provided clear evidence that these pretraining patterns were learned. In the following two sessions, the Ap patterns were paired with

additional predictive distractors, Bp, and the learning of these additional distractors was assessed relative to control patterns Cp and Dp. The test data from Sessions 3 and 4 provided no evidence for cue competition in this task. In fact, learning was greater for the Bp patterns than for the Cp and Dp patterns. Thus, these data suggest that augmented learning, rather than competition, occurred between the two sets of distractor patterns.

It is worth noting that the augmentation effect found in Experiment 2 was shown only when cuing was assessed relative to responding in the random pattern conditions. In fact, although we might have expected responses on ArBr and CrDr trials to be equivalent, responses on ArBr trials were in fact slower,  $t(19) = 2.74, p < .05$ . One potential explanation for this difference is based on the generalization occurring between predictive and random patterns. If we consider that the randomly generated ArBr patterns will share some overlap with the features of the ApBr patterns, a greater impairment in performance might be expected to occur on ArBr trials; a target is expected in a location predicted by the Ap pattern, but the target occurs elsewhere. However, it is also the case that by this explanation, the same performance impairment would be expected in the Training 2 data, and this pattern was not observed (see Figure 8B).

<sup>5</sup> Unlike in Experiment 1, the stimuli used for the C and D distractor sets are presented an equivalent number of times over the course of Experiment 2. As such, we would expect equivalent learning about Cp and Dp sets; therefore, the average of the cuing effect on these trials can be used as a comparison with Ap and Bp. A blocking effect would be shown by a smaller difference in performance between ArBp and ArBr compared with the difference between CpDr/CrDp and CrDr.

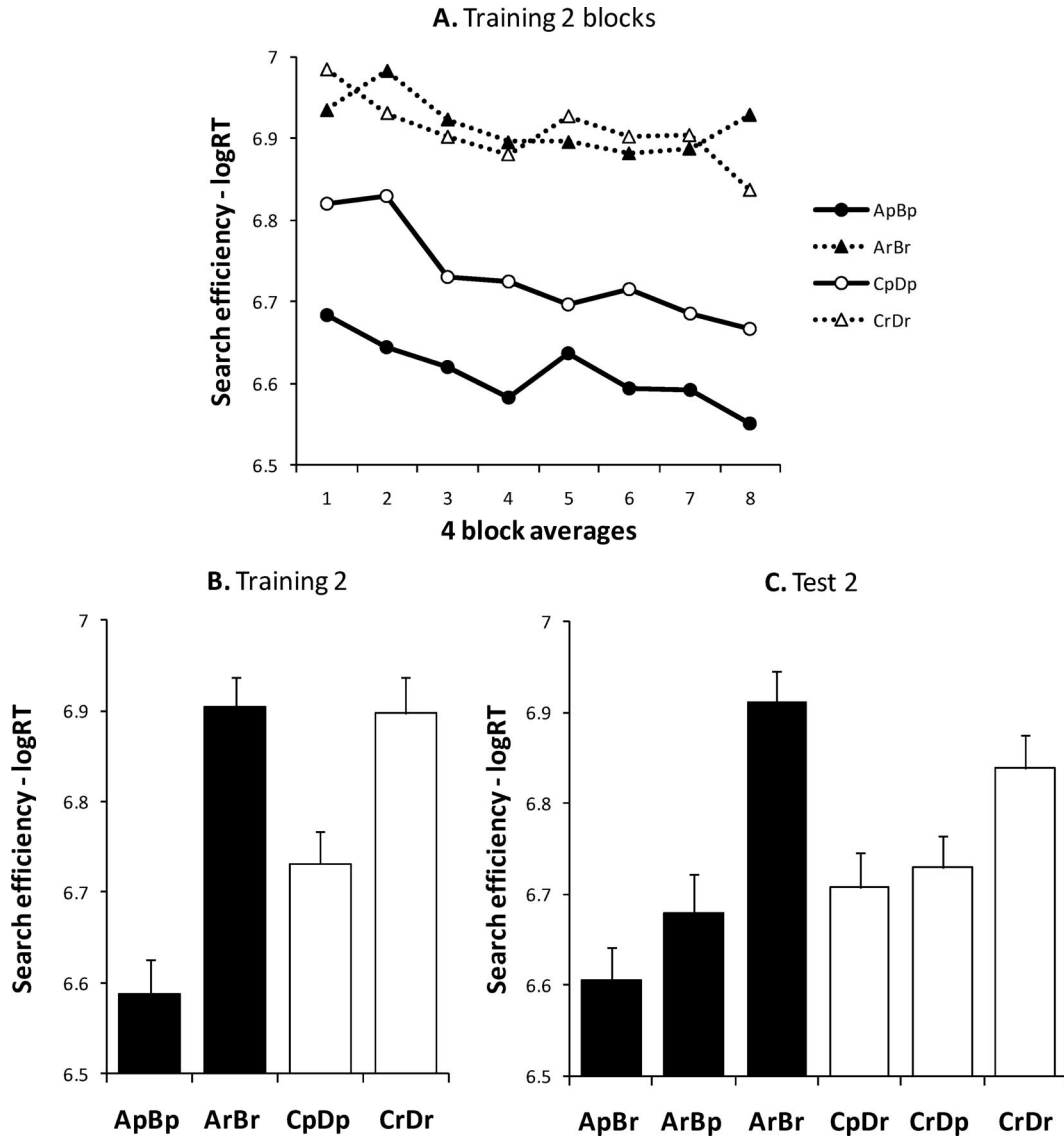


Figure 8. Target search efficiency during Sessions 3 and 4 of Experiment 2 in log transformed reaction time (logRT). Panel A shows Training 2 data collapsed into four-block averages across both sessions. Panel B shows Training 2 data averaged across all blocks and both sessions. Panel C shows Test 2 data averaged across all blocks and both sessions. Uppercase letters A through D refer to different sets of eight distractor stimuli. Lowercase letters refer to whether a set of distractors was predictive (p) of the target location or randomly arranged (r).

Alternatively, the performance impairment may reflect a general impairment on all trials presented in the AB stimulus set. That is, participants may come to expect the compound ApBp patterns on trials presented in the AB feature set. When these patterns are expected but are not presented during the test phase, participants may emit slower responses in general on all trials using the AB feature set. Thus, the baseline rate of responding might be slower in general for the AB set, and the extent to which the separate Ap and Bp patterns have been learned, or the extent of generalization between these patterns and the compound patterns, will determine whether cued responses counteract this effect on baseline responses. In contrast, we might expect the CpDp patterns to have

weaker representations because of the absence of any pretraining; therefore, we would expect this general impairment in responding for the CD set to be weaker.

### General Discussion

Propositional reasoning accounts of human learning assume that cue competition effects arise through the effortful evaluation of the contingent relationships among stimuli. A prediction that follows naturally from this account is that under conditions that do not require or allow such processes to operate, cue competition should not be observed. We explored this prediction by examining cue

competition under conditions that we hoped would restrict the operation of higher order reasoning processes and that would, in turn, promote the operation of associative learning mechanisms. Two experiments used the contextual cuing task in which learning was measured in online performance rather than offline judgments. Participants received instructions to locate and respond to a target stimulus but were given no direction toward task cues or toward the fact that learning was even possible in the task. A standard contextual cuing effect was observed: Participants were faster to detect targets when the spatial arrangement of distractor locations cued the target location. To examine cue competition under these conditions distractor–target contingencies were trained in blocking procedures. However, despite clear evidence that participants were able to learn about the contingencies between distractors and the target, there was no evidence of cue competition. These data are therefore consistent with the propositional reasoning framework (e.g., Mitchell, De Houwer & Lovibond, 2009), which predicts that cue competition will not be observed when learning occurs through incidental means.

In contrast, the absence of cue-competition effects in these data is clearly inconsistent with associative accounts that explain cue competition as resulting from a *summed error* learning process (e.g., Rescorla & Wagner, 1972). According to these accounts, cue competition is a result of the prediction error on a trial being determined by the associative strength of all the presented cues. Thus, for a blocked cue, the prediction error is determined by the combined associative strength of the pretrained and the blocked cue. Because the associative strength of the pretrained cue will lead to minimal prediction error, learning about the blocked cue will be minimal. As we have shown in the simulation results for each experiment, this prediction extends to models of contextual cuing based on similar associative mechanisms (Brady & Chun, 2007). In this respect, it is worth noting that these predictions follow for any connectionist model based on this form of competitive learning rule. For example, blocking effects are clearly predicted by models of category learning built on similar error-correction principles, and yet Bott, Hoffman, and Murphy (2007) provided data that are inconsistent with this prediction.

The extent to which the present data are also inconsistent with other models of associative learning is less clear. Models based on a *separable* rather than a summed error term (e.g., Bush & Mosteller, 1951) cannot account for cue competition. One could, therefore, appeal to such processes to explain the current data, but it is not clear why such processes would operate during incidental learning tasks and not in other forms of HCL in which cue competition has been observed. Thus, such a mechanism would be unable to provide an independent account of learning across these conditions and an appeal would have to be made to additional processes (such as inferential reasoning) to account for the data from HCL.

Perhaps a more plausible account is provided by models of associative learning, which incorporate attentional mechanisms, because these models provide an alternative explanation of cue-competition effects like blocking and overshadowing. Mackintosh (1975), for instance, suggested that learning about a cue was governed by two factors: the prediction error for that cue alone (i.e., a separable error term) and the attention paid to that cue. Here, blocking is not a direct consequence of the inability of the blocked cue to acquire associative strength through error-driven processes, because in the Mackintosh model, error-driven learning

is determined by the prediction error of each cue individually. Rather, Mackintosh proposed that because the blocked cue is a poorer predictor of the outcome than the pretrained cue on compound trials, it receives less attention, and it is this that leads to a failure to learn about blocked cues. Indeed, evidence from human contingency learning experiments has shown that attentional processes play an important role in cue-competition effects (Beesley & Le Pelley, 2011; Kruschke, Kappenman, & Hetrick, 2005; Wills, Lavric, Croft, & Hodgson, 2007). If we assume that associative learning in contextual cuing is also driven by attentional processes akin to those proposed by Mackintosh, then a failure to observe blocking in this task would point to a failure of selective attention mechanisms to tune out blocked cues.

A number of studies that have looked at selective attention mechanisms in contextual cuing speak to this issue. For example, Jiang and Chun (2001; see also Jiang & Leung, 2005) used contextual cuing tasks in which half of the distractor stimuli were presented in red, whereas the other half of the distractors were presented in green. Participants were explicitly instructed to attend to the red distractors and to ignore the green distractors. When the attended but not the unattended distractors were predictive of target location, a contextual cuing effect was observed. However, when only the unattended distractors were predictive of target location, a contextual cuing effect was not observed. These results highlight the importance of selective attention in contextual cuing, provided that a strong manipulation of attention is used (i.e., attention to color). Are these findings also informative with respect to why cue competition did not occur in the current contextual cuing tasks? In the current experiments, the different sets of distractor patterns were presented as different letter shapes (i.e., *F*, *L*, *R*, *B*). Distractor shapes presented on the same trial were perceptually similar to one another and were primarily chosen to make the main task reasonably difficult for participants. Therefore, it is likely that the perceptual grouping by object form in our experiments would not have occurred as readily as grouping by color. Moreover, in our experiments, there was no explicit instruction to attend to certain sets of distractors at the expense of others. Any biases in selective attention would therefore need to have arisen from experience with the task contingencies. Further experimental work will be required to examine whether cue competition arises when distractors are grouped along more perceptually distinct dimensions.

Although it may be possible to account for the absence of cue-competition effects in terms of a failure of selective attention mechanisms or the operation of separable error learning processes, it is difficult to see how an attentional explanation could be extended to account for the evidence for augmentation of learning in Experiment 2. One possible extension would be to assume that training with ApBr led to some overall attentional bias toward the sets of patterns presented in the stimuli comprising both the A and B patterns, such that, in general, attention was greater on AB trials than CD trials. However, if this were the case, one would expect to see facilitated responding for ArBr over CrDr trials during training. This was not the case in any phase of Experiment 2. Therefore, although one could imagine how the Brady and Chun (2007) model could be modified to incorporate such attentional mechanisms, these changes would not be expected to result in a model that generates augmentation effects. Rather, given the op-

posite nature of augmentation and blocking, it would seem that wholesale changes to the model are required.

Further support for augmentation effects can be found in a recent study by Vadillo and Matute (in press), who used a human contingency learning task in which participants were trained with a standard cue-competition procedure: A+, followed by AB+ and CD+. As we have seen, learning about cue B in such a procedure should—on the basis of error-driven learning—be blocked relative to cue D. Contrary to this prediction, Vadillo and Matute observed that when participants' decisions were made under time pressure (3 s), learning for cue B was greater than for cue D, whereas this effect was not observed when this time pressure was removed (6 s). These data support the findings presented here in suggesting that under conditions in which higher order reasoning processes are eliminated (under time pressure, incidental learning), augmentation is observed for cues trained in compound.

There exists a wealth of data on potentiation and augmentation in the animal conditioning literature, primarily in conditioned flavor aversion learning (e.g., Batsell & Batson, 1999; Bouton, Jones, McPhillips, & Swartzentruber, 1986; Rusiniak, Hankins, Garcia, & Brett, 1979). Recently, Urcelay and Miller (2009) explored the conditions under which overshadowing and potentiation occur in fear conditioning, with their results favoring a configural explanation of the potentiated learning observed for cues trained in compound. Configural theory (e.g., Pearce, 1987) states that experiencing a compound of cues leads to the formation of a configural representation of that entire stimulus compound. Cue-competition effects are accounted for by assuming a generalization decrement: When two cues are trained in compound but tested apart, the ability of each individual cue to activate the configural representation is determined by the similarity between the cue and the configural representation. For example, overshadowing arises because test cue A is less similar to its trained configural representation, AB, than C (a cue trained separately) is to its trained representation, C. To explain their potentiation effect, Urcelay and Miller appealed to low generalization decrement, such that cue A would be able to strongly activate the representation of the compound.

To what extent can the present data be explained through a modified configural theory, such as that suggested by Urcelay and Miller (2009)? Because our task was not designed specifically to encourage or discourage configural processing, it is difficult to say. However, the stimuli in this task were presented in an intermixed fashion, with an even number of distractor elements from each predictive set displayed in each quadrant of the screen. Perhaps arranging stimuli in this way, where visual cues are maximally integrated, naturally leads to a configural mode of processing (cf. Jiang & Wagner, 2004). It is possible, therefore, that some additional process that promotes augmentation of learning (such as the formation of a configural representation or strong within-compound associations) could be masking any cue-competition effects that may otherwise be observed.

It is also worth noting that the spatial integration of cues in this task resonates with those procedures used to examine cue-competition effects in spatial learning in animals. Although cue-competition effects are sometimes absent in spatial learning (see, e.g., Cheng & Newcombe, 2005), they have been observed frequently in both rats (e.g., Horne & Pearce, 2009; Roberts & Pearce, 1999) and humans (Alexander, Wilson, & Wilson, 2009).

Thus, it is unclear whether the spatial nature of the task is particularly relevant to the absence of cue competition.

A number of studies in the implicit learning literature have touched on the issue of cue competition. Endo and Takeda (2004) used a contextual cuing task in which both the distractor identities and the configuration of distractor locations cued the location of the target stimulus (Experiment 1). In this task, participants learned associations between the pattern of distractor locations and the target location but not between the pattern of distractor identities and target location. However, when distractor configurations were made irrelevant, participants did learn the associations between distractor identities and target location (Experiment 2). These data therefore show overshadowing of one cue (distractor identity) by a potentially more salient or task-relevant cue (distractor configurations).

How might one explain the contrast between the results of Endo and Takeda's (2004) study and the current data? The experiments reported here were designed to examine cue competition in the contextual cuing task using cues that were equally valid and represented by equally salient stimuli. Therefore, one explanation is that although more relevant or salient task cues might be processed more readily in the contextual cuing task (Endo & Takeda, 2004), cues of equal salience do not compete during learning.

In a recent study, Jiménez and Vázquez (2011) gave participants a contextual cuing task in which participants' responses followed a sequenced pattern. That is, participants were able to use the contextual information in the display to locate the target and were also able to prepare a response to that target on the basis of the prior sequence of responses. This dual-predictor condition was compared with ones in which participants were trained on either just the contextual information or just the sequence of responses. The results showed that learning of the context and sequence were just as strong when learned together as when learned apart, suggesting that these task cues did not compete for access to the learning mechanism.

These results clearly support the current finding that cue competition does not occur in incidental learning. However, it is worth noting a crucial dissimilarity between the designs we used and those used by Jiménez and Vázquez (2011). In their task, participants were able to learn about two separate contingencies: between the context and the target location and between the previous responses and the current response. From the perspective of error-driven learning mechanisms, the use of separate outcomes (target location, target response) is of crucial importance, because cue competition is generated by cues competing for the limited associative strength that can be accrued to a single outcome. Therefore, it is not obvious that associative models would predict a cue-competition effect in Jiménez and Vázquez's task. In contrast, in our own task, we trained cues in competition for the same outcome (target location), and as we have shown, the predictions of an associative model of contextual cuing were clear.

Finally, several studies in the domain of sequence learning have examined the effect of learning multiple task contingencies. In a typical sequence learning task, participants respond to a target stimulus that can appear in one of several locations, learning incidentally the sequenced movements of the target, such that responses are faster following predictable sequenced transitions than after random movements. Cleeremans (1997) used a version of this task in which participants were also given a highly salient cue—a cross under one

location—which signaled the location of the target on each trial. Despite the presence of this highly valid cue, sequence learning was unimpaired, suggesting that incidental learning was not subject to cue competition (see also Jiménez & Méndez, 2001; Mayr, 1996). Although we do not agree that these data necessarily support the idea of independent implicit and explicit learning systems (as has been suggested), it is clear that the lack of competition under incidental learning conditions is consistent with the idea that cue competition relies on controlled reasoning processes.

### Conclusion

The present data demonstrate that independent contextual cues of a target location do not compete for learning resources. This result is clearly inconsistent with associative learning accounts of cue competition that rely solely on error-driven learning mechanisms (Brady & Chun, 2007; Rescorla & Wagner, 1972). Although other associative learning models may, in principle, be able to account for an absence of cue competition by appealing to attentional mechanisms (e.g., Mackintosh, 1975; Pearce & Hall, 1980), the current data are clearly consistent with the idea that cue-competition effects in human learning are observed only when higher order reasoning processes are engaged (i.e., standard HCL tasks). Our data not only suggest that cue competition did not occur but that learning may even be enhanced when these cues are trained in compound. Such effects are inconsistent with the predictions of Brady and Chun's (2007) model. Future work will explore, both empirically and computationally, associative mechanisms that allow for augmentation effects in contextual cue learning.

### References

- Alexander, T., Wilson, S. P., & Wilson, P. N. (2009). Blocking of spatial learning based on shape. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 694–708.
- Batsell, W. R., Jr., & Batson, J. D. (1999). Augmentation of taste conditioning by a preconditioned odor. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*, 374–388.
- Beesley, T., & Le Pelley, M. E. (2011). The influence of blocking on overt attention and associability in human learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*, 114–120.
- Bott, L., Hoffman, A. B., & Murphy, G. L. (2007). Blocking in category learning. *Journal of Experimental Psychology: General*, *136*, 685–699.
- Bouton, M. E. (2007). *Learning and behavior: A contemporary synthesis*. Sunderland, MA: Sinauer Associates.
- Bouton, M. E., Jones, D. L., McPhillips, S. A., & Swartzentruber, D. (1986). Potentiation and overshadowing in odor-aversion learning: Role of method of odor presentation, the distal–proximal cue distinction, and the conditionability of odor. *Learning and Motivation*, *17*, 115–138.
- Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 798–815.
- Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*, 313–323.
- Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review*, *12*, 1–23.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *71*, 28–71.
- Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual 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.
- Cleeremans, A. (1997). Sequence learning in a dual-stimulus situation. *Psychological Research*, *60*, 72–86.
- De Houwer, J., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, *55*, 965–985.
- De Houwer, J., Beckers, T., & Vandorpe, S. (2005). Evidence for the role of higher order reasoning processes in cue competition and other learning phenomena. *Learning & Behavior*, *33*, 239–249.
- Dickinson, A., Shanks, D. R., & Evenden, J. L. (1984). Judgement of act–outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology: A. Human Experimental Psychology*, *36*, 29–50.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Perception & Psychophysics*, *66*, 293–302.
- Fletcher, P. C., Anderson, J. M., Shanks, D. R., Honey, R., Carpenter, T. A., Donovan, T., . . . Bullmore, E. T. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nature Neuroscience*, *4*, 1043–1048.
- Goujon, A., Didierjean, A., & Marmèche, E. (2009). Semantic contextual cuing and visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 50–71.
- Hoffmann, J., & Sebald, A. (2005). Local contextual cuing in visual search. *Experimental Psychology*, *52*, 31–38.
- Horne, M. R., & Pearce, J. M. (2009). A landmark blocks searching for a hidden platform in an environment with a distinctive shape after extended pretraining. *Learning & Behavior*, *37*, 167–178.
- Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology*, *54*, 1105–1124.
- Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review*, *12*, 100–106.
- Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—Configuration or individual locations? *Perception & Psychophysics*, *66*, 454–463.
- Jiménez, L., & Méndez, C. (2001). Implicit sequence learning with competing explicit cues. *Quarterly Journal of Experimental Psychology: A. Human Experimental Psychology*, *54*, 345–369.
- Jiménez, L., & Vázquez, G. A. (2011). Implicit sequence learning and contextual cueing do not compete for central cognitive resources. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 222–235.
- Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York, NY: Appleton-Century-Crofts.
- Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 830–845.
- Kunar, M. A., Flusberg, S. J., Horowitz, T. S., & Wolfe, J. M. (2007). Does contextual cueing guide the deployment of attention? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 816–828.
- Le Pelley, M. E., Beesley, T., & Suret, M. B. (2007). Blocking of human causal learning involves learned changes in stimulus processing. *Quarterly Journal of Experimental Psychology*, *60*, 1468–1476.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298.
- Macphail, E. M. (1982). *Brain and intelligence in vertebrates*. Oxford, United Kingdom: Clarendon.



- Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 350–364.
- Miller, R. R., Barnet, R., & Grahame, N. (1995). Assessment of the Rescorla–Wagner model. *Psychological Bulletin*, 117, 363–386.
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, 32, 183–198.
- Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, 9, 273–302.
- Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, 94, 61–73.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87, 532–552.
- Rausei, V., Makovski, T., & Jiang, Y. (2007). Attention dependency in implicit learning of repeated search context. *Quarterly Journal of Experimental Psychology*, 60, 1321–1328.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Roberts, A. D. L., & Pearce, J. M. (1999). Blocking in the Morris swimming pool. *Journal of Experimental Psychology: Animal Behavior Processes*, 25, 225–235.
- Rowland, L. A., & Shanks, D. R. (2006). Attention modulates the learning of multiple contingencies. *Psychonomic Bulletin & Review*, 13, 643–648.
- Rusiniak, K. W., Hankins, W. G., Garcia, J., & Brett, L. P. (1979). Flavor–illness aversions: Potentiation of odor by taste in rats. *Behavioral & Neural Biology*, 25, 1–17.
- Schultz, W., & Dickinson, A. (2000). Neural coding of prediction errors. *Annual Review of Neuroscience*, 23, 473–500.
- Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology*, 61, 273–301.
- Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory & Cognition*, 36, 403–415.
- Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, 88, 135–170.
- Urcelay, G. P., & Miller, R. R. (2009). Potentiation and overshadowing in Pavlovian fear conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 35, 340–356.
- Vadillo, M. A., & Matute, H. (in press). Augmentation in contingency learning under time pressure. *British Journal of Psychology*.
- van Asselen, M., & Castelo-Branco, M. (2009). The role of peripheral vision in implicit contextual cuing. *Attention, Perception, & Psychophysics*, 71, 76–81.
- Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning. *Nature*, 412, 43–48.
- Wills, A. J., Lavric, A., Croft, G. S., & Hodgson, T. L. (2007). Predictive learning, prediction errors, and attention: Evidence from event-related potentials and eye tracking. *Journal of Cognitive Neuroscience*, 19, 843–854.

## Appendix

### Technical Details of the Brady and Chun (2007) Model

Two sets of weights connect input and output units. The fixed weights impose spatial constraints reflecting an *attentional spotlight* surrounding the target, such that the model learns more about distractors that occur nearest to the targets on repeating trials. The fixed weights are derived from an exponential function of the distances between stimuli:

$$FW_{oi} = S \times e^{-d_{oi}}, \quad (\text{A1})$$

where  $FW$  is the fixed weight,  $d_{oi}$  is the number of matrix locations between input unit  $i$  and output unit  $o$ , and  $S$  is a constant that determines the diameter of the attentional spotlight.

The learned weights represent the associations between the stimulus positions and the target locations. A large weight between input unit  $i$  and output unit  $o$  reflects the model's prediction that the presence of a stimulus at location  $i$  predicts the occurrence of the target stimulus at location  $o$ .

On each trial, the pattern of activation on the input units is fed through the weights to produce a pattern of activation on the output units:

$$in_o = \left( \sum a_i \times LW_{oi} \times FW_{oi} \right) + \beta_o, \quad (\text{A2})$$

$$a_o = \frac{1}{1 + e^{-in_o}}, \quad (\text{A3})$$

where  $in_o$  is the summed activation passed to output unit  $o$  from the input units;  $a_i$  is the activation of input unit  $i$ ;  $LW_{oi}$  and  $FW_{oi}$  are the learned weight (or association) and the fixed weight, respectively, between input unit  $i$  and output unit  $o$ ;  $\beta_o$  is a bottom-up activation term, which ensures higher activation values to output units representing locations in which stimuli are present;  $a_o$  is the activation of output unit  $o$ .

The learned weights are adjusted using a form of the delta rule:

$$\delta_o = \lambda - a_o, \quad (\text{A4})$$

$$(\Delta LW_{oi})_t = \varphi \times \delta_o \times a_o \times a_i + \theta \times (\Delta LW_{oi})_{t-1}, \quad (\text{A5})$$

where  $\delta_o$  is the prediction error of output unit  $o$ , determined by the difference between the expected value of the output unit,  $\lambda$ , and the actual activation of output unit  $o$ ,  $a_o$ . In Equation A5,  $\varphi$  is a learning rate parameter determining the contribution of the current prediction error to the weight change,  $a_i$  is the activation of input unit  $i$ , and  $\theta$  is a momentum term that determines how much the previous weight change ( $t - 1$ ) contributes to the direction and magnitude of the current change ( $t$ ).

Received June 15, 2010

Revision received May 16, 2011

Accepted June 3, 2011 ■