Journal of Experimental Psychology: Learning, Memory, and Cognition

Learned Predictiveness Influences Rapid Attentional Capture: Evidence From the Dot Probe Task

Mike E. Le Pelley, Miguel Vadillo, and David Luque

Learned Predictiveness Influences Rapid Attentional Capture: Evidence From the Dot Probe Task

Mike E. Le Pelley
University of New South Wales and Cardiff University

Miguel Vadillo
University College London

David Luque
University of Málaga

Attentional theories of associative learning and categorization propose that learning about the predictiveness of a stimulus influences the amount of attention that is paid to that stimulus. Three experiments tested this idea by looking at the extent to which stimuli that had previously been experienced as predictive or nonpredictive in a categorization task were able to capture attention in a dot probe task. Consistent with certain attentional theories of learning, responses to the dot probe were faster when it appeared in a location cued by a predictive stimulus compared to a location cued by a nonpredictive stimulus. This result was obtained only with short (250-ms or 350-ms) but not long (1,000-ms) delays between onset of the stimuli and the dot probe, suggesting that the observed spatial cuing effect reflects the operation of a relatively rapid, automatic process. These findings are consistent with the approach to the relationship between attention and learning taken by the class of models exemplified by Mackintosh's (1975) theory.

Keywords: associative learning, attention, categorization, dot probe task

Supplemental materials: http://dx.doi.org/10.1037/a0033700.supp

Attention provides the gateway between the mass of information in the world and the relatively small subset of that information that we select for further analysis or action. What, then, determines the stimuli that will be selected under a given set of circumstances? Much of the research related to this issue that is described in the cognitive psychology literature has focused on the "intrinsic" perceptual and emotional properties of stimuli. For example, a stimulus is more likely to capture attention if it is highly perceptually salient (e.g., if it has an abrupt onset, or a bright color; Folk, Remington, & Johnston, 1992) or if it is "emotionally relevant": Negative mood states bias attention toward threatening information (MacLeod, Mathews, & Tata, 1986), while positive mood states bias attention to desirable, rewarding stimuli (Tamir & Robinson, 2007). A second line of research has instead examined the ability of external events to modulate attention to stimuli. For example, the influence of attention to a stimulus persists for longer if selection of that stimulus is highly rewarded than if it is only weakly rewarded (Della Libera & Chelazzi, 2006).

While the studies described above have considered the effects of stimulus and reward properties on attentional selection in isolation, recent research suggests that we should also consider them in combination (Anderson, Laurent, & Yantis, 2011; Della Libera & Chelazzi, 2009; Hickey, Chelazzi, & Theeuwes, 2010; Kiss, Driver, & Eimer, 2009; Le Pelley, Mitchell, & Johnson, 2013). For example, Della Libera and Chelazzi (2009) gave participants training on a task in which selection of certain shapes was typically followed by high reward (€0.10), while selection of other shapes was typically followed by low reward (€0.01). After extensive training, shapes that predicted high-value outcomes were shown to be easier to select when serving as targets (Experiment 2), and more difficult to reject when serving as distractors (Experiment 1), compared to shapes that predicted low-value outcomes. Using a visual search task, Anderson et al. (2011) similarly demonstrated that presenting cues previously associated with high-value reward as distractors led to a slowing of search, compared to cues previously associated with low-value reward. These findings suggest that cues associated with high-value outcomes are more likely to capture attention than those paired with low-value outcomes; that is, participants learn to attend to cues as a function of the value of the reward with which they are paired. Consequently these studies can be described as demonstrating an influence of *learned value* on attention.

Other studies of associative learning in both humans and nonhuman animals suggest that the learned *predictiveness* of a stimulus might also be a determinant of attention to that stimulus (for

reviews, see Le Pelley, 2004; Mitchell & Le Pelley, 2010). The predictiveness of a stimulus refers to the accuracy with which the occurrence of that stimulus allows subsequent events to be predicted. A predictive stimulus is one that is consistently followed by the same outcome (which can be of high or low value, such that predictiveness and learned value are orthogonal); a nonpredictive stimulus is not. In a typical experiment, participants learn during Phase 1 that certain stimuli are predictive of outcomes, while others are not.[1] In a subsequent Phase 2 involving new stimulus–outcome contingencies, human participants typically learn faster about stimuli previously experienced as predictive than those previously experienced as nonpredictive (e.g., Bonardi, Graham, Hall, & Mitchell, 2005; Kruschke, 1996; Le Pelley & McLaren, 2003; Le Pelley, Suret, & Beesley, 2009; Le Pelley, Turnbull, Reimers, & Knipe, 2010). Such findings support the suggestion that attention is modulated by learned predictiveness, as long as it is assumed that differences in the rate of learning about stimuli during Phase 2 reflect differences in attention to those stimuli, as suggested by "attentional" theories of associative learning (e.g., Kruschke, 2003; Le Pelley, 2004; Mackintosh, 1975). However, this assumption is questionable. After all, the putative relationship between attention and rate of learning has been invoked only to account for the results of experiments of this kind and has received little external validation in the cognitive psychology literature. Consequently, the possibility remains open that what is influenced by predictiveness, and what in turn influences learning, is not attention but rather an *associability* parameter that merely modulates the rate at which stimuli enter into associations; a "nonattentional" model along these lines has been proposed by Honey, Close, and Lin (2010; see also Oswald et al., 2001). Alternatively, one could appeal to memory processes, rather than attention. Perhaps stimuli experienced as predictive during Phase 1 develop stronger and/or more distinct representations in memory than those experienced as nonpredictive, and this allows information experienced in Phase 2 to be more accurately addressed to (associated with) the correct stimulus representation for predictive stimuli than nonpredictive stimuli (see Le Pelley, Reimers, et al., 2010). Yet another nonattentional account of these rate-of-learning studies has recently been proposed, in terms of an "inferential–attribution" process (Mitchell, Griffiths, Seetoo, & Lovibond, 2012).

A growing set of studies has taken a more direct approach to investigating the putative relationship between learned predictiveness and attention, by looking at the effect of predictiveness on measures of attention that have previously been validated in the cognitive psychology literature. In particular, predictiveness has been shown to influence overt attention, measured in terms of gaze location (Beesley & Le Pelley, 2011; Kruschke, Kappenman, & Hetrick, 2005; Le Pelley, Beesley, & Griffiths, 2011; Wills, Lavric, Croft, & Hodgson, 2007). However, shifting attention does not necessarily entail eye movements. Visual performance can be enhanced at the site where attention is directed without changing fixation (Jonides, 1981; Yeshurun & Carrasco, 1999). Hence, it is desirable to develop a test of the influence of learned predictiveness on attention that does not rely on eye gaze. Moreover, all of these previous gaze-based studies have demonstrated a bias in overt attention at the point at which participants made their categorization response. It is perhaps unsurprising that participants are more likely to pay attention to a stimulus during a categorization task if identification of that stimulus is necessary for making a

correct categorization response. A more powerful finding would be a demonstration that learning about the predictiveness of stimuli produces a more general attentional bias with regard to those stimuli that operates even when it is not required, and when it may even hinder performance. Finally, and more pragmatically, eye tracking apparatus is typically expensive, intrusive and cumbersome. Consequently, it is not well-suited for use with large participant samples, or for testing outside the laboratory—for example, with children in schools, or with patients in in-patient facilities. This last point is pertinent, given that a dysfunction of the relationship between learning and attention has been implicated in schizophrenia (Morris, Griffiths, Le Pelley & Weickert, 2012), Parkinson's disease (Gauntlett-Gilbert, Roberts, & Brown, 1999; Hampshire & Owen, 2010), obsessive-compulsive disorder (Hampshire & Owen, 2010), as well as certain types of brain injury (Owen, Roberts, Polkey, Sahakian, & Robbins, 1991). It would therefore be advantageous to develop a procedure to measure the relationship between attention and learning that can be implemented on any standard computer without special equipment and can be used with multiple participants simultaneously.

Toward this end, the current experiments examined the influence of predictiveness on attentional capture using a variant of the spatial cuing task (Posner, Nissen, & Ogden, 1978). It is well established that responses to targets appearing in an attended location are faster than to targets appearing in an unattended location (Posner, 1980; Posner, Snyder, & Davidson, 1980). Following this rationale, Folk et al. (1992) demonstrated that responses to a target were faster when it appeared in a location previously occupied by a cue defined in terms of an abrupt onset or a discontinuity in color. Similarly, on each trial of MacLeod et al.'s (1986) *dot probe task*, a pair of words appeared—one threat-related (e.g., *injury*) and the other neutral. Anxious participants were faster to identify a target (a small dot) when it subsequently appeared in the location that had been occupied by the threat-related word compared to the neutral word. The implication of these studies is that stimuli defined in terms of abrupt onsets, color discontinuities, or emotional valence, can capture attentional resources.

This type of cuing procedure is the original and classic paradigm used for investigating the operation of attentional processes in humans but has never before been used to assess the relationship between attention and learned predictiveness. In the current experiments, we use this approach to investigate whether stimuli that differ in predictiveness also differ in the extent to which they capture spatial attention, and in Experiments 2 and 3 we go on to look at the timecourse over which this capture occurs.

Experiment 1

Experiment 1 involved two tasks. The first was described to participants as a categorization task, with the design shown in

---

[1] Notably, in such studies all stimuli are typically followed by equally valued outcomes on all trials and, hence, do not differ in terms of their learned *value*. The stimuli differ only in terms of the accuracy with which they predict which specific outcome will occur, i.e., in terms of their learned *predictiveness*. This is also the case in the experiments described in the current article.

Table 1. On each trial, two stimuli appeared: a square filled with one of two shades of green (labeled Gre1 and Gre2), and a set of oblique lines at one of two orientations (Lin1 and Lin2; see Figure S1 in the online supplemental materials). Participants categorized this pair of stimuli into one of two categories by making an appropriate response, with immediate corrective feedback provided. For participants in condition "Green Predictive," the shade of green predicted the correct categorization response and the orientation of the lines was nonpredictive: Presence of Gre1 indicated that response R1 was correct, presence of Gre2 indicated that response R2 was correct, while Lin1 and Lin2 provided no information on the correct response. For participants in condition "Lines Predictive" the orientation of the lines was predictive of the correct response and shade of green was nonpredictive. If predictiveness influences attention, then this categorization task should cause participants to come to attend more strongly to the stimuli belonging to the predictive dimension than to those belonging to the nonpredictive dimension.

The second task was designed to assess any such difference in attention, using a variant of the dot probe procedure. On each trial one of the green squares and one of the sets of oblique lines appeared briefly on either side of the screen. A dot probe target could then appear in the location previously occupied by one of these cues. This target was equally likely to appear in the location cued by the stimulus that had been predictive during the categorization task as the location cued by the stimulus that had been nonpredictive. Hence, both stimuli were equally valid as cues during the dot probe task. If the stimulus that was predictive during the categorization task was more likely to capture attention, however, then responses to the target should be faster when it appeared in the location cued by this stimulus, compared to the location cued by the nonpredictive stimulus.

These two tasks alternated across phases—Experiment 1 contained four task phases, in the order: categorization, dot probe, categorization, dot probe. This procedure meant that learning about the categorization predictiveness of cues was "topped up" prior to each iteration of a relatively short test on the dot probe task.

Table 1
Design of the Categorization Task

Stimulus pair	Correct response	
	Green predictive	*Lines predictive*
Gre1 & Lin1	R1	R1
Gre1 & Lin2	R1	R2
Gre2 & Lin1	R2	R1
Gre2 & Lin2	R2	R2

Note. Gre1 and Gre2 refer to squares filled with slightly different shades of green; Lin1 and Lin2 refer to oblique lines at slightly different orientations (Experiment1) or lines with different thicknesses (Experiment 2). Particular values on each dimension were randomly assigned to these labels for each participant. R1 and R2 refer to two distinct responses. The correct response for each stimulus pair is shown for participants in the "green predictive" condition (in which the shade of green was predictive of the correct response while the orientation of the oblique lines was nonpredictive) and the "lines predictive" condition (in which the orientation [Experiment 1] or thickness [Experiment 2] of the oblique lines was predictive of the correct response while the shade of green was nonpredictive).

Method

Participants and apparatus. Eleven Cardiff University students (10 female) took part in exchange for course credit and were tested individually in a dim, quiet room. Stimuli were presented on a 43.2-cm monitor, and stimulus presentation was controlled by a Visual Basic program. Timing used Windows API QueryPerformance functions for millisecond resolution. Responses were made using the keyboard, and error signals were beeps given over headphones.

Stimuli. The two green squares had red–green–blue color values of (0, 255, 0) and (0, 160, 0), with sides subtending 3.42° visual angle from a viewing distance of 60 cm. The two oblique line stimuli comprised sets of cyan lines (thickness 0.34° visual angle) sloping upward to the right at an angle of either 33° or 57°, enclosed within a black square background with sides 3.42° visual angle. Stimuli are shown in Figure S1.

These green square and oblique line stimuli were presented centrally in white square frames with sides subtending 3.76°, which were positioned either side of a small fixation cross located in the center of the screen; the distance from the center of the cross to the center of each box subtended 5.30°. The target in the dot probe task was a white equilateral triangle with side length subtending 2.22°. This would appear centrally in one of the white square frames. The screen background was black.

Design. Two between-subjects conditions were created by varying the predictive dimension during the categorization task. Participants were initially assigned randomly to conditions, and following the exclusions described below replacements were run to ensure equal numbers in each condition. Overall, four participants were tested in condition Green Predictive and seven in condition Lines Predictive. Particular values on each stimulus dimension were randomly assigned to the labels shown in Table 1 for each participant; e.g., for some participants in condition Lines Predictive, the label Lin1 in Table 1 referred to lines at 33° to the horizontal and Lin2 referred to lines at 57°, while for others this assignment was reversed.

Experiment 1 contained four task phases, in the order: categorization task, dot probe task, categorization task, dot probe task. Each phase of the categorization task was split into blocks. Each block contained four trials, with each of the stimulus pairs shown in Table 1 appearing once in random order. The first phase of the categorization task had 12 blocks, and the second phase had eight blocks. Across blocks, for each stimulus pair the predictive stimulus appeared equally often on the left and on the right.

On each trial of the dot probe task, one of the four stimulus pairs shown in Table 1 appeared as the cue. This pair could appear with the green square on the left or on the right. On target-present trials, the target could subsequently appear on the left or the right. For target-present trials, every combination of stimulus pair, stimulus position, and target position appeared once during each phase of the dot probe task (giving $4 \times 2 \times 2 = 16$ target-present trials). Each phase of the dot probe task also included eight target-absent trials; one for each combination of stimulus pair and stimulus position. The 24 trials of each phase of the dot probe task were presented in random order.

Procedure. Participants received written and oral descriptions of the tasks. Initial instructions described the categorization task. Participants were told that (a) on each trial a cross would appear in

the center of the screen and that they should keep their eyes fixed on it throughout the trial; (b) a pair of stimuli would then be presented; (c) their task was to decide if that pair of stimuli belonged to Category 1 (in which case they should press the C key) or to Category 2 (M key); (d) they would start out guessing but that on the basis of feedback their decisions should become more accurate; (e) they should make as few errors as possible; (f) they should respond within 3 s but should not anticipate the stimuli. There followed four practice trials using stimuli that were not encountered in the main body of the experiment and with no feedback provided. Following these practice trials, participants were asked if they had maintained central fixation throughout each trial; if they had not, the practice trials were repeated.

Participants then received instructions relating to the dot probe task as follows: (a) a cross would appear in the center of the screen and that it was very important that their eyes remained fixed on it; (b) two stimuli would be presented briefly; (c) after these stimuli disappeared, the target might appear to the left or right of the central cross; (d) if this target appeared, they should press the spacebar, and if it did not appear, they should do nothing; (e) they should respond as quickly as possible but should not anticipate the target. There followed four practice trials of this task (three target-present and one target-absent), using stimuli not encountered in the body of the experiment. If participants reported not maintaining fixation on these trials, they were repeated.

Participants were then shown the four stimuli to be used in the body of the experiment, before the experiment began.

Each trial of the categorization task began with the appearance of the fixation cross flanked by the two empty white frames. After 1,000 ms, the stimulus pair appeared in these frames. If participants made the correct category response, the word "Correct" appeared in the center of the screen; if they made the incorrect response, "Wrong" appeared. If a response was made within 150 ms, the message "Do not anticipate the stimuli" appeared. If participants did not make a response within 3 s, the message "You took too long" appeared. All non-"correct" feedback was accompanied by an error signal. Feedback remained on screen for 800 ms; the screen then cleared and the next trial began after 600 ms. After every eight trials of the categorization task participants were told how many errors they had made in those eight trials, and their mean response time (RT; excluding any anticipations or timeouts). This information remained on-screen for 3 s before the experiment proceeded.

After the first phase of the categorization task, participants read a reminder of the instructions for the dot probe task. Figure 1 shows a schematic of a target-present trial of this task. Each trial began with presentation of the fixation cross flanked by two white frames. After 1,000 ms, the stimulus pair was presented for 150 ms in these frames, and then disappeared. On target-present trials, the dot probe appeared after a delay of 200 ms, giving a stimulus-onset asynchrony (SOA) of 350 ms between presentation of the stimulus pair and the dot probe. If participants pressed the spacebar within 2,000 ms, the screen cleared and the next trial began after an interval of 600 ms. If participants did not press the spacebar within 2,000 ms on target-present trials, a timeout error signal was given and the experiment moved to the next trial. Responses within 150 ms of the dot probe appearing were deemed anticipations and an error signal was given. On target-absent trials, no dot probe appeared; if participants pressed the space-
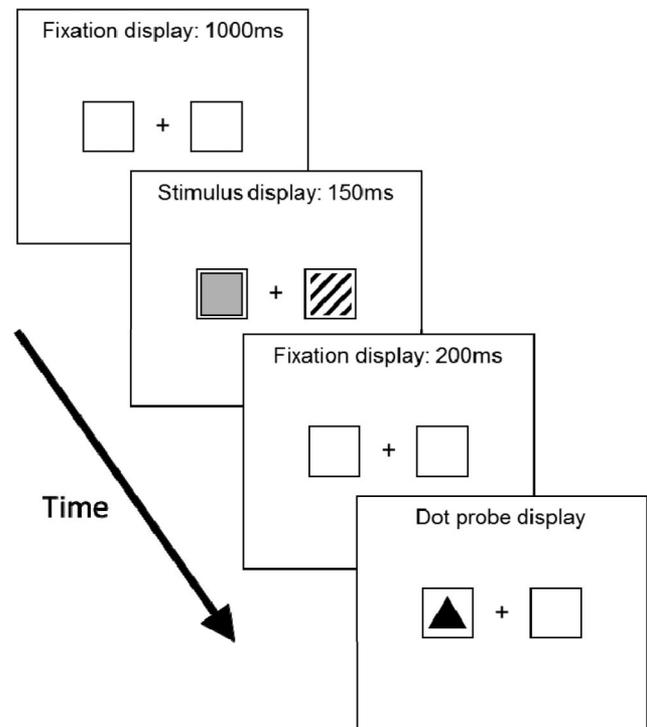

Figure 1. Schematic example of the sequence of events on a target-present trial of the dot probe task in Experiment 1. The gray square in the cue display represents one of the green squares that could be used as cue stimuli, and the striped square represents one of the sets of oblique lines.

bar during a 2,150-ms window after the cues vanished this was deemed an anticipation, and an error signal was given. If they did not respond, then after this window had elapsed the next trial proceeded. After 16 trials of the dot probe task participants were told how many anticipations they had made during those trials, and mean RT on target-present trials.

The procedure for the second categorization and dot probe phases was the same. Instructions prior to each phase stated which task participants would be carrying out in that phase, and reminded them of the importance of maintaining fixation on each trial.

Results

An influence of predictiveness on spatial cuing in the dot probe task could only be expected if participants managed to learn about the differential predictiveness of the cues involved during the categorization task. Following Le Pelley and McLaren (2003), a selection criterion of 60% correct categorization responses averaged across all blocks of the categorization task was imposed (chance performance = 50% correct). Three participants (all in condition Lines Predictive) failed to achieve this criterion, and their data were excluded from all subsequent analyses.

The primary focus of this article is the influence of predictiveness on attentional capture, as opposed to any "unlearned" influence of stimulus salience (in terms of color, intensity, onset, etc.). In other words, we are interested in the status of cues as predictive or nonpredictive, rather than in the particular identity of these cues (green squares or sets of oblique lines). To reflect this focus—and

since, following the exclusions described above, the number of participants in each condition was equal—the two between-subjects conditions (Green Predictive and Lines Predictive) were combined. In this pooled sample, the label *predictive stimuli* refers to green squares for participants from condition Green Predictive, and sets of oblique lines for participants from condition Lines Predictive; *nonpredictive stimuli* refers to oblique lines for participants from condition Green Predictive, and green squares for participants from condition Lines Predictive. (See supplemental materials for an analysis including condition [Green Predictive vs. Lines Predictive] as a between-subjects factor.)

Figure 2 shows accuracy across training blocks of the categorization task. One-way analysis of variance (ANOVA) revealed a significant effect of block, $F(19, 133) = 2.82$, $\eta_p^2 = .29$, $p < .001$, with accuracy increasing across training (effect size in this and all subsequent analyses is partial eta-squared, η_p^2). Collapsing across the blocks of each phase (Phase 1 being Blocks 1–12 and Phase 2 being Blocks 13–20) revealed that mean accuracy was significantly greater than chance (50%) in both Phase 1 and Phase 2, one-sample $t(7) = 3.39$ and 13.7, $\eta_p^2 = .62$ and $.96$, $p = .012$ and $p < .001$, respectively.

Accuracy on the dot probe task was very high; across all participants, only one anticipation was made on a target-absent trial. The data of main interest relate to the target-present trials. Anticipation responses and timeouts (which constituted 1.2% and 0% of all target-present trials, respectively) were removed. Several measures were taken to reduce the impact of any outlying response times (RTs). First, RTs were log-transformed. Second, any log RTs lying more than three standard deviations from each participant's mean were excluded as outliers (1.95% of all target-present trials), following Sinich (1986).

Target-present trials of the dot probe task were labeled *congruent* if the dot probe appeared in the location cued by the predictive stimulus from the categorization task, and *incongruent* if the dot probe appeared in the location cued by the nonpredictive stimulus. For each participant we calculated the median RT for congruent

and incongruent trials for each phase of the dot probe task; these data were then averaged across participants and are shown in Figure 3. ANOVA with factors of congruence and phase revealed a main effect of congruence, $F(1, 7) = 15.3$, $\eta_p^2 = .69$, $p = .006$, indicating that across all participants, responses were significantly faster on congruent trials than incongruent trials. The observed difference in log RTs between congruent and incongruent trials corresponds to an RT difference of 14 ms. The main effect of phase and the phase × congruence interaction were nonsignificant, larger $F(1, 7) = 0.74$, both $\eta_p^2 s < .1$, both $p s > .40$.

Discussion

Participants were faster to respond to the appearance of the dot probe when it appeared in a location that was cued by a stimulus that had been predictive in the categorization task, relative to a stimulus that had been nonpredictive. The implication of this finding is that the difference in experienced predictiveness of the stimuli influenced their tendency to capture attention during the dot probe task. This relationship between predictiveness and attention is exactly that anticipated by attentional theories of associative learning (e.g., Kruschke, 2003; Le Pelley, 2004; Mackintosh, 1975).

Notably, the attentional capture observed in Experiment 1 occurred even though there was no advantage to be gained in shifting attention to the predictive stimulus during the dot probe task. That is, because the dot probe was equally likely to appear in the location cued by the predictive stimulus as that cued by the nonpredictive stimulus, the best strategy during this task was to maintain central fixation throughout (and participants were explicitly instructed to do so). The implication, then, is that the faster responses to congruent than incongruent dot probes observed in Experiment 1 might reflect relatively automatic shifts of attention toward stimuli experienced as being predictive during the categorization task, or away from those experienced as nonpredictive. This kind of automatic, stimulus-driven or exogenous attentional orienting can be contrasted with endogenous shifts of attention that are under the control of the participant (Jonides, 1981; Posner, 1980; Posner & Cohen, 1984).

Experiment 2

The aims of Experiment 2 were twofold. The first aim was to replicate our novel finding of faster responding in the dot probe task when the probe appears in a location congruent with a stimulus that has been experienced as predictive of a categorization response. The second aim relates to the issue of automaticity, raised in the Discussion of Experiment 1. We argued there that during the dot probe task, there was no reason for participants to consciously shift their attention toward one stimulus rather than the other, and no advantage to be gained in doing so. Nevertheless it remains possible that they did so regardless, strategically orienting attention toward the predictive stimulus for some reason. If this is the case, and the dot probe data of Experiment 1 reflect a conscious strategy of shifting attention toward predictive stimuli, then providing more time to process the predictive status of the stimuli (in terms of a longer SOA between stimuli and probe in the dot probe task) should produce stronger or at least similar effects, assuming that such controlled strategies are time-consuming (see
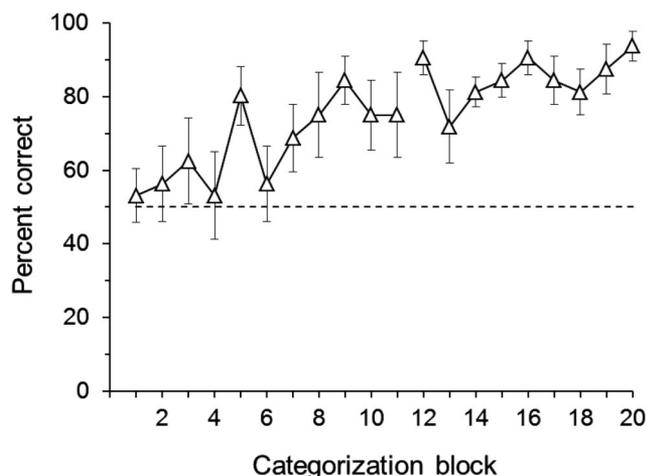
Figure 2. Mean percentage correct categorization responses across the 20 blocks of Experiment 1. Error bars show standard error of the mean. Blocks 1–12 constituted Phase 1, and Blocks 13–20 constituted Phase 2. Dotted line shows level of performance expected by chance (50% correct).
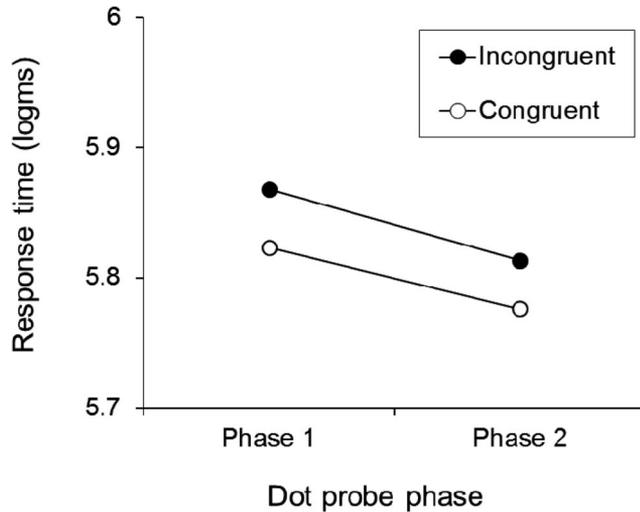
Figure 3. Median log-transformed response time to appearance of the probe on target-present trials of the dot probe task of Experiment 1. Congruent = probe appeared in location cued by predictive stimulus from categorization task; Incongruent = probe appeared in location cued by nonpredictive stimulus from categorization task.

De Houwer, Hermans, & Eelen, 1998; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Le Pelley, Calvini, & Spears, 2013; Posner & Snyder, 1975). Experiment 2 investigated this issue by varying the SOA in the dot probe task as a within-subject variable.

Method

Participants and apparatus. Seventy-two University of Málaga students (47 female) participated in exchange for course credit. They were tested in groups of up to 10 at a time in a room containing 10 semienclosed cubicles, using standard PCs with 48.3 cm monitors. Stimulus presentation was controlled by the Cogent 2000 toolbox (http://www.vislab.ucl.ac.uk/Cogent/) for MATLAB. Participants made all responses with the computer keyboard using their dominant hand.

Stimuli. The two green squares had red–green–blue color values of (51, 255, 0) and (0, 102, 0), with sides subtending 9.0° visual angle from a distance of 60 cm. The two oblique line stimuli comprised sets of thick (width = 0.86° visual angle) or thin (width = 0.01°) rightward-slanted blue lines, enclosed within a black square background with sides subtending 9.0°. Stimuli are shown in Figure S2 (see online supplemental materials).

These green square and oblique line stimuli were presented centrally in white square frames with sides subtending 13.1°, which were positioned either side of a small fixation cross that was located in the center of the screen; the distance from the center of the cross to the center of each box subtended 9.0°. The dot probe was a white square with side length subtending 1.35°. This would appear superimposed centrally on one of the stimuli. The screen background was black.

Design. For half of the participants the shade of the green square determined the correct response in the categorization task (condition Green Predictive); for the other half, the thickness of the oblique lines determined the correct response (condition Lines

Predictive). Particular values on each stimulus dimension (shade of green and thickness of lines) were randomly assigned to the labels shown in Table 1 for each participant.

Procedure. Initial instructions (in Spanish) described the categorization task. Participants were told that on each trial a pair of stimuli would appear and that they were required to make a response using either the up or down arrow keys and that their task was to learn the correct response for each stimulus pair. Participants then completed a first phase of 32 categorization trials. This phase comprised four, eight-trial blocks, with each of the stimulus pairs in Table 1 appearing twice per block in random order; for each stimulus pair, the predictive stimulus appeared once on the left and once on the right. Incorrect responses produced the feedback message "Error! The correct response was [UP/DOWN]," which remained onscreen for 3 s; correct responses were not followed by any explicit feedback.

Participants then moved on to the first phase of the dot probe task. Instructions for this task were similar to those of Experiment 1, but participants were now told explicitly that, in order to respond to the square (the dot probe target) as quickly as possible, "it is best to ignore the figures" (i.e., the stimulus pair). Each dot probe trial began with presentation of a central fixation cross. After 500 ms the stimulus pair appeared to either side of this cross. After an SOA of either 250 ms or 1,000 ms, the dot probe appeared superimposed on one of the stimuli. This probe remained until participants made the correct response (left arrow key for a target presented on the left; right arrow key for a target presented on the right). Immediately on making the correct dot probe response, the screen cleared, and the next trial began after an intertrial interval of 1 s.

Each phase of the dot probe task contained 16 trials: 2 SOAs (250 ms or 1,000 ms) × 4 stimulus pairs (see Table 1) × 2 trial types (target congruent with predictive stimulus versus target incongruent with predictive stimulus). Whether the predictive stimulus appeared on the left or right was randomly determined on each trial.

After the first phase of the dot probe task, participants returned to the categorization task. Experiment 2 comprised eight alternations of the categorization task with the dot probe task.

Results

Three participants failed to achieve the criterion of 60% correct averaged over all the trials of the categorization task (two in condition Green Predictive and one in condition Lines Predictive). These participants' data were excluded from further analysis. As for Experiment 1, data were collapsed across counterbalancing conditions Green Predictive and Lines Predictive; see supplemental materials for an analysis including condition (Green Predictive vs. Lines Predictive) as a between-subjects factor.

Figure 4 shows accuracy across training blocks of the categorization task. One-way ANOVA revealed a significant effect of block, $F(31, 2108) = 65.3$, $\eta_p^2 = .49$, $p < .001$, with accuracy increasing across training. Collapsing across the blocks of each phase revealed that mean accuracy was significantly greater than chance (50%) in all phases, smallest $t(68) = 11.3$, all $\eta_p^2 s > .65$, all $ps < .001$.

Dot probe trials were defined as correct if participants' first response correctly corresponded to the position of the probe.
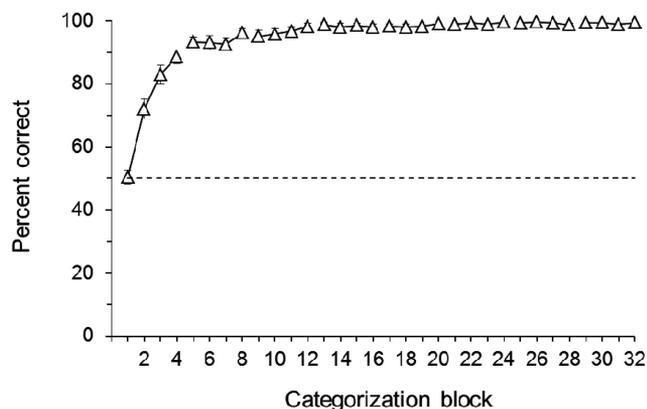
Figure 4. Mean percent correct categorization responses across the 32 blocks of Experiment 2. Error bars show standard error of the mean. The level of performance expected by chance is 50% correct. Each four blocks constituted one phase of the categorization task (i.e., Blocks 1–4 constituted Phase 1, Blocks 5–8 constituted Phase 2, etc.).

Accuracy on the dot probe task was very high, with a mean of 99.3 ± 0.15% (SEM) correct trials across all participants.

In Experiment 1, the experiment program defined minimum and maximum RT limits of 150 ms and 2,000 ms for the dot probe task, as responses faster or slower than these limits, respectively, were not permitted. For consistency, dot probe RTs in Experiment 2 that were below 150 ms or above 2,000 ms were therefore excluded from analysis (0.24% and 0.01% of all trials, respectively). As for the analysis of Experiment 1, RTs were log-transformed and any log RTs lying more than three standard deviations from each participant's mean were excluded as outliers (1.46% of all trials). As in Experiment 1, for each participant we calculated the median log RT for congruent and incongruent trials for each phase of the dot probe task (using data from correct trials only). These data were analyzed as four epochs, with each epoch representing the averaged data from a consecutive pair of dot probe phases (see Figure 5).

At an SOA of 250 ms, a cuing effect occurred with faster responses to the probe on congruent trials than incongruent trials across all epochs. In contrast, at an SOA of 1,000 ms there was no clear cuing effect; RTs on congruent and incongruent trials were, on average, more similar. A $(2) \times (2) \times (4)$ repeated-measures ANOVA with factors of SOA, congruence and epoch revealed no significant main effect of congruence, $F(1, 68) = 1.71$, $\eta_p^2 = .02$, $p = .20$. Crucially, however, congruence interacted with SOA, $F(1, 68) = 4.17$, $\eta_p^2 = .06$, $p = .045$, indicating that the influence of congruence was significantly greater at 250-ms SOA than at 1,000-ms. The main effect of SOA was significant, $F(1, 94) = 1.85$, $\eta_p^2 = .73$, $p < .001$, with faster responses on trials with 1,000-ms SOA (presumably because the longer SOA allowed more time for response preparation). The main effect of epoch was significant, $F(3, 204) = 2.83$, $\eta_p^2 = .04$, $p = .040$, although Figure 5 reveals no continuing pattern in changes of RT across epochs. Other interactions were nonsignificant, largest $F(3, 204) = 1.04$, $\eta_p^2 s < .012$, $p s > .38$.

Analysis of simple effects (collapsing across epochs) revealed that the effect of congruence at 250-ms SOA was significant, $F(1, 68) = 4.81$, $\eta_p^2 = .07$, $p = .032$, corresponding to an advantage of

5 ms for congruent trials over incongruent trials. There was no effect of congruence at 1,000-ms SOA, $F(1, 68) = 0.16$, $\eta_p^2 = .002$, $p = .69$.

Discussion

Experiment 2 replicated the cuing effect observed in Experiment 1 (faster responses on congruent than incongruent trials of the dot probe task), but only for trials with 250-ms SOA. Trials with 1,000-ms SOA showed no difference in response times when the dot probe appeared in the location of the stimulus that was predictive in the categorization task, relative to when it appeared in the location of the nonpredictive stimulus.

In the Introduction to Experiment 2 we argued that, if the cuing effect observed at short SOA reflected a conscious strategy of shifting attention toward predictive stimuli, then providing more time to process the predictive status of the stimuli should produce stronger or at least similar effects. In contrast, the data reveal that providing more time led to a significant *weakening* of the cuing effect. The implication is that the cuing effect observed at short SOA is not a consequence of a controlled, strategic process, but rather reflects an automatic process. This issue is taken up in the General Discussion.

Experiment 3

Attentional theories of associative learning (e.g., Kruschke, 2003; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall, 1980)
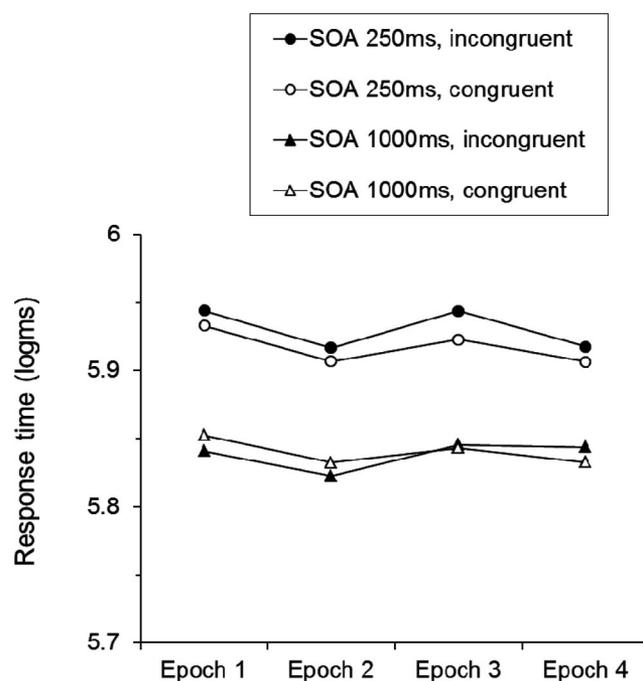

Figure 5. Median log-transformed response time to the dot probe in Experiment 2 for trials with a stimulus-onset asynchrony (SOA) of 250 ms or 1,000 ms. Congruent = probe appeared in location cued by predictive stimulus from categorization task; Incongruent = probe appeared in location cued by nonpredictive stimulus from categorization task. Each of Epochs 1–4 comprises two consecutive phases of the dot probe task (i.e., Phases 1 and 2 constitute Epoch 1, Phases 3 and 4 constitute Epoch 2, etc.).

suggest that changes in attention to stimuli develop as a consequence of learning about the predictive status of those stimuli; that is, an attentional bias should develop incrementally as a function of how well participants have learned the various stimulus–response associations. Experiments 1 and 2 did not provide clear support for this prediction, in that the size of the attentional bias in the dot probe task remained similar across the course of each experiment (see Figures 3 and 5). However, this may be because in each case participants had received considerable training on the categorization task prior to the first phase of the dot probe task, such that the predictive status of the stimuli was already quite well established at this point. Figures 2 and 4 show that categorization accuracy in the final block of the first phase of the categorization task had reached a similar level to that maintained subsequently.

Experiment 3 used a procedure that combined the categorization and dot probe task. This allowed us to assess more closely the development of the dot probe bias across the course of training, and thus provided a more thorough test of the suggestion that the size of the attentional bias should increase as function of learning about the predictive status of the stimuli.

Combining the categorization and dot probe phases brings a further advantage. While the cuing effect at short SOA observed in Experiments 1 and 2 was significant, it was numerically small. This may be because any automatic attentional bias toward the predictive cue that is learned during the categorization task extinguishes to some extent over the trials of the following dot probe phase. If this was indeed the case, then combining the two tasks such that every dot probe trial is also a categorization trial would minimize the influence of extinction and, hence, should result in a larger cuing effect at short SOA.

Experiment 3 also incorporated the SOA manipulation of Experiment 2, although this was now varied between-subjects rather than within subject. Given the findings of Experiment 2, we expected to observe a cuing effect for participants with short SOA in the dot probe task, but no effect for participants with long SOA.

Method

Participants, apparatus and stimuli. A total of 108 University of Málaga students (86 female) participated in exchange for course credit. Testing conditions, apparatus and stimuli were as for Experiment 2.

Design. In Experiment 3, the SOA of the dot probe task was varied between subjects. For 53 participants, the SOA between presentation of the stimuli and the dot probe was 250 ms; of these, 27 participants were in condition Green Predictive, and 26 were in condition Lines Predictive. For the remaining 55 participants (27 in condition Green Predictive and 28 in condition Lines Predictive) the SOA was 1,000 ms. Particular values on each stimulus dimension (shade of green and thickness of lines) were randomly assigned to the labels shown in Table 1 for each participant.

Procedure. Initial instructions for the categorization task were as in Experiment 2. After reading these instructions, participants completed a pretraining phase of 16 categorization trials, with each of the stimulus pairs in Table 1 appearing four times in random order; for each stimulus pair, the predictive stimulus appeared twice on the left and twice on the right. Categorization feedback was as for Experiment 2.

Following this pretraining phase, participants received further instructions explaining that subsequent trials would be more complicated: On each trial (a) a pair of stimuli would appear; (b) a small white square (the dot probe) would then appear superimposed on one of these stimuli; (c) participants should press the left arrow if the square appeared on the left stimulus, and the right arrow if it appeared on the right; (d) once they had responded to the square, they should make a response to the stimulus pair using the up or down arrows as in the pretraining stage. Similar to Experiment 2, participants were told that they should respond to the position of the dot probe as rapidly as possible and that "In order to do so, it is best that you ignore the pair of figures until you have responded to the location of the square." Finally, participants were told that occasionally an arrow would appear in the center of the screen and that when it did they should press the corresponding arrow key as rapidly as possible. These "arrow trials" were intended to further encourage participants to maintain central fixation at the start of each trial, since unexpected targets could occasionally appear at this location.

Figure 6 shows a schematic of a standard trial. Each such trial began with presentation of a central fixation cross. After 500 ms the stimulus pair appeared to either side of this cross. After an SOA of either 250 ms or 1,000 ms (depending on which between-subjects condition the participant had been allocated to), the dot probe appeared superimposed on one of the stimuli. This probe remained until participants made the correct response (left arrow key for a target
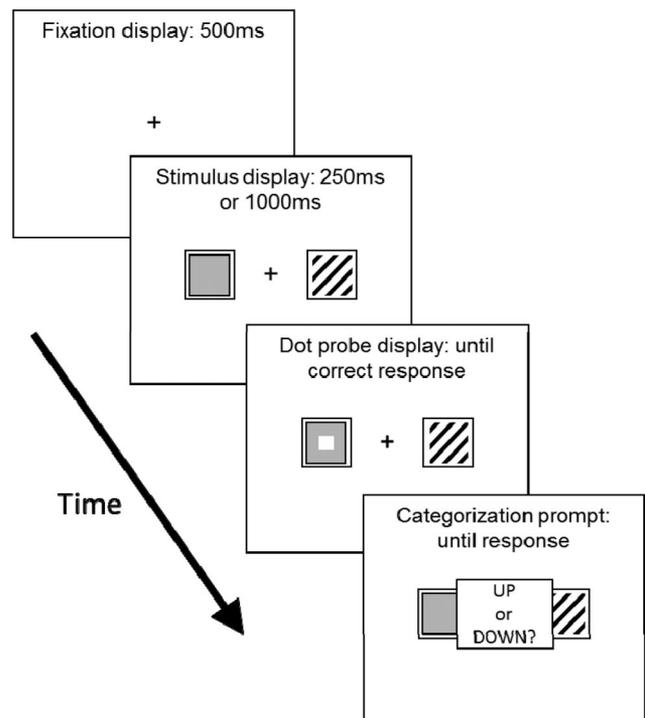

Figure 6. Schematic example of the sequence of events on a standard trial of Experiment 3. The gray square in the cue display represents one of the green squares that could be used as cue stimuli, and the striped square represents one of the sets of oblique lines. The duration of the stimulus display (and, hence, the stimulus-onset asynchrony of the dot probe task) was varied between subjects.

presented on the left; right arrow key for a target presented on the right). Immediately on making the correct dot probe response, the probe disappeared and the message "UP or DOWN?" appeared. Participants then made a categorization response using the up or down arrow keys; feedback was administered as during the pretraining phase, and the next trial then began after an intertrial interval of 1 s.

On "arrow trials," the fixation cross appeared for 500 ms and was then replaced by a small arrow in the center of the screen, which remained until participants made a response using one of the arrow keys. No feedback was provided for these responses. The next trial then began after an interval of 1 s.

Participants completed 15 blocks of trials. Each block contained 17 trials in random order—one arrow trial (with the direction of the arrow determined randomly), and every combination of the four trial types in Table 1 with the predictive stimulus appearing on the left or the right, and with the dot probe appearing on the left or the right ($4 \times 2 \times 2 = 16$ trials). Hence, as in previous experiments, the dot probe was equally likely to appear in the location of the predictive or the nonpredictive stimulus.

Results

Twelve participants failed to achieve the selection criterion of 60% correct when averaged over all the trials of the categorization task (five in condition Green Predictive and two in condition Lines Predictive with SOA = 250 ms; three in condition Green Predictive and two in condition Lines Predictive with SOA = 1,000 ms). These participants' data were excluded from further analysis. As for Experiments 1 and 2, data were collapsed across counterbalancing conditions Green Predictive and Lines Predictive; see supplemental materials for an analysis including condition (Green Predictive vs. Lines Predictive) as a between-subjects factor. Data were analyzed as five epochs, with each epoch representing the averaged data from three training blocks.

Figure 7 shows accuracy of participants' categorization responses. Accuracy rose steadily across the course of training and was similar in participants who experienced an SOA of 250 ms or 1,000 ms in the dot probe task. A $2 \times (5)$ ANOVA with factors of
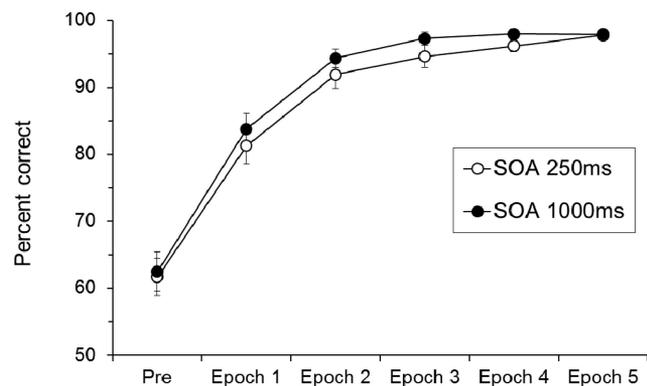
SOA and epoch revealed a significant main effect of epoch, $F(4, 376) = 50.9$, $\eta_p^2 = .35$, $p < .001$, but no main effect of SOA, $F(1, 94) = 1.58$, $\eta_p^2 = .02$, $p = .21$, or interaction, $F(4, 376) = 0.34$, $\eta_p^2 = .004$, $p = .85$. (Note that that this analysis did not include the data from the pretraining phase because trials in this phase did not incorporate the dot probe task, and we are interested in how participants' accuracy relates to their dot probe performance.)

Accuracy on the dot probe task was very high, with a mean of $99.3 \pm 0.10\%$ (standard error of the mean) correct trials across all participants. Dot probe RT data were processed as for Experiment 2. RTs below 150 ms or above 2,000 ms were excluded (0.06% and 0.92% of all trials, respectively). RTs were then log-transformed and any log RTs lying more than three standard deviations from each participant's mean were excluded as outliers (0.63% of all trials). For each participant we calculated the median log RT for congruent and incongruent trials for each block of training (using data from correct trials only). These data were then averaged across epochs of three blocks; see Figure 8.

At an SOA of 250 ms, a clear cuing effect occurred with faster responses to the probe on congruent trials than incongruent trials, and the size of this effect increased across epochs. In contrast, at an SOA of 1,000 ms there was no cuing effect; RTs on congruent and incongruent trials remained similar to one another across all epochs but showed a general decrease as the experiment progressed. A $2 \times (2) \times (5)$ ANOVA with factors of SOA, congruence and epoch revealed a significant main effect of congruence, $F(1, 94) = 25.2$, $\eta_p^2 = .21$, $p < .001$. Crucially, the effect of congruence interacted with SOA, $F(1, 94) = 20.5$, $\eta_p^2 = .18$, $p < .001$, indicating that the influence of congruence was significantly greater at 250 ms SOA than at 1,000 ms. Congruence also interacted with epoch, $F(4, 376) = 3.67$, $\eta_p^2 = .04$, $p = .006$, with the influence of congruence tending to increase across epochs. The main effect of epoch was significant, $F(4, 376) = 5.08$, $\eta_p^2 = .05$, $p < .001$, with RTs decreasing across epochs as participants became more familiar with the task. The SOA $\times$ Epoch interaction was also significant, $F(4, 376) = 6.63$, $\eta_p^2 = .07$, $p < .001$, with the decrease in RT across epochs being greater for participants with 1,000 ms SOA than for those with 250 ms. The main effect of SOA was significant, $F(1, 94) = 13.6$, $\eta_p^2 = .13$, $p < .001$, with faster responses in participants with 1,000-ms SOA (presumably because the longer SOA allowed more time for response preparation). The three-way interaction did not reach significance, $F(4, 376) = 1.68$, $\eta_p^2 = .02$, $p = .15$.

This omnibus analysis was followed up by separate, preplanned two-way ANOVAs using the data from each group of participants defined by SOA. For participants with 250-ms SOA, there was a significant main effect of congruence, $F(1, 45) = 32.2$, $\eta_p^2 = .42$, $p < .001$, with shorter RTs on congruent than incongruent trials. The observed mean difference in log RTs between congruent and incongruent trials corresponds to an RT difference of 46 ms. The Congruence $\times$ Epoch interaction was significant, $F(4, 180) = 5.20$, $\eta_p^2 = .10$, $p < .001$, with the size of the congruence effect tending to increase across epochs, showing a significant linear trend, $F(1, 45) = 14.8$, $\eta_p^2 = .25$, $p < .001$. Simple effects analysis revealed that the effect of congruence was significant in every epoch, smallest $F(1, 45) = 4.57$, all $\eta_p^2 s > .09$, all $ps < .038$. The main effect of epoch was not significant, $F(4, 180) = 1.62$, $\eta_p^2 = .03$, $p = .17$. For participants with 1,000-ms SOA, there was no main effect of congruence, $F(1, 49) = 0.19$, $\eta_p^2 = .004$, $p = .67$,

Figure 7. Mean percent correct categorization responses across Experiment 3 for participants who experienced a stimulus-onset asynchrony (SOA) of 250 ms or 1,000 ms in the dot probe task. Pre = pretraining phase; each of Epochs 1–5 comprises three blocks of 16 trials. Error bars show standard error of the mean. The level of performance expected by chance is 50% correct.
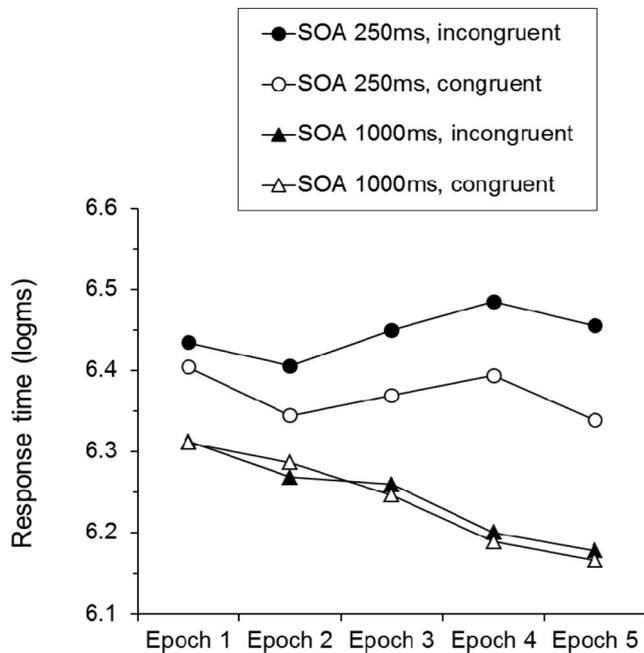
Figure 8. Median log-transformed response time to the dot probe in Experiment 3 for participants who experienced a stimulus-onset asynchrony (SOA) of 250 ms or 1,000 ms in this task. Congruent = probe appeared in location cued by predictive stimulus from categorization task; Incongruent = probe appeared in location cued by nonpredictive stimulus from categorization task. Each of Epochs 1–5 comprises three blocks of 16 trials.

and no Congruence × Epoch interaction, $F(4, 196) = 0.69$, $\eta_p^2 = .07$, $p = .60$. The main effect of epoch was significant for these participants, $F(4, 196) = 10.6$, $\eta_p^2 = .37$, $p < .001$.

The analysis described above demonstrated that the cuing effect at 250-ms SOA increased across epochs. A further analysis tested the more specific question of whether this effect increased as a function of accuracy on the categorization task. For each participant in the 250 ms SOA group, we calculated the Spearman's rank correlation between categorization accuracy and the size of the cuing effect (incongruent RT minus congruent RT) across epochs. One participant was excluded from this analysis as their categorization accuracy was equal in all epochs, such that no correlation could be calculated. The mean correlation across remaining participants was $\rho(5) = .20$, with a standard error of .07. A one-sample t test revealed that this mean correlation was significantly greater than zero, $t(44) = 2.67$, $\eta_p^2 = .14$, $p = .010$, demonstrating that the cuing effect was indeed correlated with categorization performance.

Finally, performance on the occasional "arrow trials" was similar in both SOA groups. Mean accuracy on arrow trials was $97.8 \pm 1.48\%$ for the 250 ms SOA group, and $99.5 \pm .26\%$ for the 1,000-ms SOA group; this difference was not significant, $t(95) = 1.14$, $\eta_p^2 = .01$, $p = .26$. Mean RT on arrow trials was 667 ± 16 ms for the 250-ms SOA group, and 683 ± 14 ms for the 1,000-ms SOA group; this difference was not significant, $t(95) = 0.77$, $\eta_p^2 = .006$, $p = .44$.

Discussion

Experiment 3 replicated the key findings of Experiment 2: a significant cuing effect in the dot probe task at 250-ms SOA but no cuing effect at 1,000-ms SOA. Moreover, the cuing effect at 250 ms was significantly greater than that at 1,000 ms SOA. And combining the categorization and dot probe tasks produced a cuing effect at short SOA that was numerically larger than that observed in Experiments 1 and 2 (46 ms, compared to 14 ms in Experiments 1 and 5 ms in Experiment 2). This is consistent with the suggestion that the effects in our earlier studies were smaller due to extinction of previously learned attentional biases during the dot probe task.

Importantly, in Experiment 3 the cuing effect at short SOA increased over the course of training, and more specifically increased as a function of the accuracy of participants' categorization responses. This finding is consistent with the suggestion made by attentional theories of associative learning (e.g., Kruschke, 2003; Le Pelley, 2004; Mackintosh, 1975) that learning about the predictive status of stimuli drives changes in attention to those stimuli.

General Discussion

Three experiments used a dot probe task to demonstrate that learning about the predictiveness of stimuli influences attentional orienting to those stimuli. In all experiments, when a short SOA was used in the dot probe task (350 ms in Experiment 1; 250 ms in Experiments 2 and 3), participants were faster to respond to the dot probe when it appeared in the same location as a stimulus that had previously been experienced as predictive in a categorization task, than when it appeared in the location of a stimulus experienced as nonpredictive. Experiment 3 demonstrated that this cuing effect increased in magnitude as performance on the categorization task improved.

The cuing effect occurred even though the short SOA meant that participants had little time to shift attention to the location of the predictive stimulus, and even though there was no advantage to be gained in so doing. Indeed, in Experiments 2 and 3 participants were explicitly informed that the best strategy was to ignore the stimulus pair until after they had responded to the dot probe, and unpredictable "arrow trials" were included in Experiment 3 to encourage participants to try to maintain attention to the central fixation point during the dot probe task. The fact that a cuing effect was still observed under these conditions suggests that the source of this effect was a relatively rapid, automatic process.

This suggestion is further supported by the finding of Experiments 2 and 3 that increasing the SOA to 1,000 ms did not increase or even maintain the cuing effect, as might be expected if this effect reflected a conscious strategy of shifting attention toward predictive stimuli (see De Houwer et al., 1998; Fazio et al., 1986; Le Pelley, Calvini, & Spears, 2013; Posner & Snyder, 1975). Instead, increasing SOA led to a significant reduction in the cuing effect (and in fact eliminated it entirely). It is true that in both Experiments 2 and 3 there was a difference in baseline response time, with faster responses at long SOA than short SOA, raising the possibility that the failure to observe a cuing effect at long SOA might simply be a consequence of a floor effect in response times. However, this seems unlikely. First, response times at 1,000-ms SOA decreased significantly across epochs in Experiment 3. Hence, it is clear that response times

were not at floor in the earlier epochs, and yet there was no hint of a congruency effect at 1,000-ms SOA in these epochs (while there was a significant effect at 250-ms SOA). Second, mean response time for the dot probe task at 1,000-ms SOA in Experiment 3 was *longer* than at 250-ms SOA in Experiment 2 (6.24 log ms vs. 5.92 log ms, or 514 ms vs 374 ms; see Figures 8 and 5). And yet a congruence effect was observed for the short response times at 250-ms SOA in Experiment 2 but not for the longer response times at 1,000-ms SOA in Experiment 3. Hence, it seems unlikely that the failure to observe a congruence effect at 1,000-ms SOA in Experiment 3 reflects a lack of sensitivity at this level of baseline response time.

Instead, we suggest that the cuing effect at short SOA reflects the operation of an automatic attentional process. That is, presentation of the stimuli leads to a rapid, automatic orienting of attention to the predictive stimulus,[2] producing the cuing effect at short SOA. One possibility that can account for the current data is that, during the longer SOA, this initial automatic attentional influence decays and, hence, attention returns to the center of the display, such that no cuing effect is seen at long SOA (cf. Fazio et al., 1986). However, our favored account appeals instead to an interaction between automatic and controlled attentional processes (cf. Klauer, Roßnagel, & Musch, 1997). Participants in the current task knew that the best strategy was to attend centrally during the dot probe task. On this account, the long SOA provides sufficient time for participants to use controlled processes to correct for and overcome the automatic attentional orienting caused by presentation of the stimuli, returning attention to the center of the display. This latter account has the advantage that it is also able to account for the persistence of greater attention to predictive cues that is typically observed during categorization training using eye tracking (e.g., Le Pelley et al., 2011). During categorization training in these studies there is no particular reason for participants to try to maintain central fixation. Consequently—and unlike in the dot probe task of the current study—there is no drive for them to use controlled processes to overcome an initial automatic tendency to attend to the predictive stimuli, and, hence, this tendency will persist over a longer period.

Attentional theories of associative learning (e.g., Kruschke, 2003; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall, 1980) share the central dogma that learning about the predictiveness of stimuli influences attention to those stimuli. As noted in the introduction, the majority of studies that have been conducted to test this idea have used the rate of learning about stimuli as a proxy measure of attention. This approach, however, leaves such studies open to interpretation in nonattentional terms, for example in terms of differences in the strength of mnemonic representations. By using a more direct measure of attentional orienting, the current studies provide a more direct test—and confirmation—of this central dogma of attentional theories. These data add to research that has used eye gaze as a measure of the influence of learning about predictiveness on overt attention (Beesley & Le Pelley, 2010; Kruschke et al., 2005; Le Pelley et al., 2011; Wills et al., 2007). However, the current dot probe procedure has certain important advantages over these previous studies using eye gaze (see also Livesey, Harris, & Harris, 2009). In particular, it meets the criteria laid down in the Introduction. First, we have used the dot probe procedure to demonstrate the incremental development of an attentional bias in a task (dot probe) that is incidental to the task that produces that bias (categorization); indeed, there is no need for participants to pay attention to the stimuli at all during the dot probe task. Second, the dot probe procedure does not rely on movements of eye fixation, which may or may not accompany shifts of attention (Jonides, 1981; Yeshurun & Carrasco, 1999). Third, it does not require expensive, intrusive or cumbersome equipment and can be implemented on any standard computer (including laptops or tablets that could be used outside the laboratory) and used with multiple participants simultaneously. Beyond these criteria, the dot probe procedure also allowed us to test the automaticity of the attentional bias produced by learned predictiveness by varying the SOA; studies of eye gaze have not so far provided evidence relating to this issue.

Discriminating Between Attentional Theories of Associative Learning

Up to this point, "attentional theories" of learning have been treated as a single generic class of model, but, in fact, different attentional theories take a rather different view of the relationship between predictiveness and attention. The current experiments were introduced from the perspective of the class of attentional theory exemplified by Mackintosh's (1975) model (see also Kruschke, 2003), which proposes that more attention is devoted to cues that are more accurate predictors of the current outcome (here category membership) than those that are less predictive. Clearly models taking this approach are well-equipped to account for the congruence effect observed in the current experiments.

However, an alternative attentional account was proposed by Pearce and Hall (1980), who suggested that more attention will be devoted to stimuli that are followed by surprising outcomes, than to stimuli that are followed by well-predicted and hence unsurprising outcomes. This account is less successful when applied to the current data. In its original formulation, this model states that what is crucial for determining attention is not how surprising the outcome is given the presence of a particular, individual stimulus, but rather how surprising the outcome is given the combination (compound) of all currently presented stimuli. In the categorization design used here (Table 1), all compounds are equally predictive of category membership (e.g., the compound "Gre1 & Lin2" belongs to the same category on all training trials), and hence the outcome occurring on each trial is equally surprising. Consequently, Pearce and Hall's model predicts that all stimuli will maintain equal attention throughout the experiment, which is at odds with the dot probe data of the current experiments.

---

[2] Or, potentially, an automatic shift of attention *away from* the nonpredictive stimulus. The current experiments do not allow us to decide between an account in which participants learn to shift attention toward predictive stimuli, or away from nonpredictive stimuli (or both). That said, one might expect that the cuing effect produced by a shift away from the nonpredictive stimulus would be very weak, since attention could potentially shift away from this stimulus in any direction. In other words, a shift away from the nonpredictive stimulus will not necessarily be a shift toward the location of the predictive stimulus (where the dot probe target will appear).

Certain studies of the influence of predictiveness on the rate of novel learning about a stimulus, conducted in nonhuman animals, are consistent with the approach suggested by Pearce and Hall (1980), and inconsistent with Mackintosh's (1975) model (see Le Pelley, 2004, for a review). However, studies with humans that are able to decide between these models have generally produced results consistent with Mackintosh's account (e.g., Bonardi et al., 2005; Kruschke, 1996; Le Pelley et al., 2011; Le Pelley & McLaren, 2003; Le Pelley, Turnbull, et al., 2010; see Le Pelley, 2010, for a review), and the current data add to this collection.

Conclusion

When considering the factors that influence attention, in addition to thinking about intrinsic properties of the stimulus (e.g., its size, color, onset, or emotional valence), and the external factors of rewarding attention to a particular stimulus, it would seem that we also need to consider stimulus and reward properties in combination. That is, learning of an association between a particular stimulus and a particular reward (or outcome, or category membership) will also influence the attention that is devoted to that stimulus in future. The current studies demonstrate that learning about the predictiveness of stimuli produces a relatively automatic attentional bias with respect to those stimuli; a bias that is consistent with the pattern anticipated by some of the most influential attentional theories of associative learning.

References

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 108,* 10367–10371. doi:10.1073/pnas.1104047108

Beesley, T., & Le Pelley, M. E. (2010). The effect of predictive history on the learning of sub-sequence contingencies. *The Quarterly Journal of Experimental Psychology, 63,* 108–135. doi:10.1080/17470210902831767

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. doi:10.1037/a0019526

Bonardi, C., Graham, S., Hall, G., & Mitchell, C. J. (2005). Acquired distinctiveness and equivalence in human discrimination learning: Evidence for an attentional process. *Psychonomic Bulletin & Review, 12,* 88–92. doi:10.3758/BF03196351

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective and identity priming with episodically associated stimuli. *Cognition and Emotion, 12,* 145–169. doi:10.1080/026999398379691

Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science, 17,* 222–227. doi:10.1111/j.1467-9280.2006.01689.x

Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science, 20,* 778–784. doi:10.1111/j.1467-9280.2009.02360.x

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229–238. doi:10.1037/0022-3514.50.2.229

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030–1044. doi:10.1037/0096-1523.18.4.1030

Gauntlett-Gilbert, J., Roberts, R. C., & Brown, V. J. (1999). Mechanisms underlying attentional set-shifting in Parkinson's disease. *Neuropsychologia, 37,* 605–616. doi:10.1016/S0028-3932(98)00049-9

Hampshire, A., & Owen, A. M. (2010). Clinical studies of attention and learning. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 385–406). Oxford, England: Oxford University Press.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *The Journal of Neuroscience, 30,* 11096–11103. doi:10.1523/JNEUROSCI.1026-10.2010

Honey, R. C., Close, J., & Lin, E. (2010). Acquired distinctiveness and equivalence: A synthesis. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 159–186). Oxford, England: Oxford University Press.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.

Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science, 20,* 245–251. doi:10.1111/j.1467-9280.2009.02281.x

Klauer, K. C., Roßnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 246–255. doi:10.1037/0278-7393.23.1.246

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science, 8,* 225–248. doi:10.1080/095400996116893

Kruschke, J. K. (2003). Attention in learning. *Current Directions in Psychological Science, 12,* 171–175. doi:10.1111/1467-8721.01254

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. doi:10.1037/0278-7393.31.5.830

Le Pelley, M. E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 57,* 193–243.

Le Pelley, M. E. (2010). Attention and human associative learning. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 187–215). Oxford, England: Oxford University Press.

Le Pelley, M. E., Beesley, T., & Griffiths, O. (2011). Overt attention and predictiveness in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37,* 220–229. doi:10.1037/a0021384

Le Pelley, M. E., Calvini, G., & Spears, R. (2013). Learned predictiveness influences automatic evaluations in human contingency learning. *The Quarterly Journal of Experimental Psychology, 66,* 217–228. doi:10.1080/17470218.2012.708760

Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 56,* 68–79.

Le Pelley, M. E., Mitchell, C. J., & Johnson, A. M. (2013). Outcome value influences attentional biases in human associative learning: Dissociable effects of training and of instruction. *Journal of Experimental Psychology: Animal Behavior Processes, 39,* 39–55. doi:10.1037/a0031230

Le Pelley, M. E., Reimers, S. J., Calvini, G., Spears, R., Beesley, T., & Murphy, R. A. (2010). Stereotype formation: Biased by association. *Journal of Experimental Psychology: General, 139,* 138–161. doi:10.1037/a0018210

Le Pelley, M. E., Suret, M. B., & Beesley, T. (2009). Learned predictiveness effects in humans: A function of learning, performance, or both?

Journal of Experimental Psychology: Animal Behavior Processes, 35, 312–327. doi:10.1037/a0014315

Le Pelley, M. E., Turnbull, M. N., Reimers, S. J., & Knipe, R. L. (2010). Learned predictiveness effects following single-cue training in humans. *Learning & Behavior, 38,* 126–144. doi:10.3758/LB.38.2.126

Livesey, E. J., Harris, I. M., & Harris, J. A. (2009). Attentional changes during implicit learning: Signal validity protects a target stimulus from the attentional blink. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 408–422. doi:10.1037/a0014525

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276–298. doi:10.1037/h0076778

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95,* 15–20. doi:10.1037/0021-843X.95.1.15

Mitchell, C. J., Griffiths, O., Seetoo, J., & Lovibond, P. F. (2012). Attentional mechanisms in learned predictiveness. *Journal of Experimental Psychology: Animal Behavior Processes, 38,* 191–202. doi:10.1037/a0027385

Mitchell, C. J., & Le Pelley, M. E. (Eds.). (2010). *Attention and associative learning: From brain to behaviour.* Oxford, England: Oxford University Press.

Morris, R. W., Griffiths, O., Le Pelley, M. E., & Weickert, T. W. (2012). Attention to irrelevant cues is related to positive symptoms in schizophrenia. *Schizophrenia Bulletin.* Advance online publication. doi:10.1093/schbul/sbr192

Oswald, C. J. P., Yee, B. K., Rawlins, J. N. P., Bannerman, D. B., Good, M., & Honey, R. C. (2001). Involvement of the entorhinal cortex in a process of attentional modulation: Evidence from a novel variant of an IDS/EDS procedure. *Behavioral Neuroscience, 115,* 841–849. doi:10.1037/0735-7044.115.4.841

Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., & Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal-lobe excisions, temporal-lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia, 29,* 993–1006. doi:10.1016/0028-3932(91)90063-E

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian conditioning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552. doi:10.1037/0033-295X.87.6.532

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32,* 3–25. doi:10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X* (pp. 531–556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick & I. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale, NJ: Erlbaum.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solo (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

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. doi:10.1037/0096-3445.109.2.160

Sincich, T. (1986). *Business statistics by example.* San Francisco, CA: Dellen.

Tamir, M., & Robinson, M. D. (2007). The happy spotlight: Positive mood and selective attention to rewarding information. *Personality and Social Psychology Bulletin, 33,* 1124–1136. doi:10.1177/0146167207301030

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. doi:10.1162/jocn.2007.19.5.843

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research, 39,* 293–306. doi:10.1016/S0042-6989(98)00114-X