Journal of Experimental Psychology:
Animal Behavior Processes
2011, Vol. 37, No. 2, 220–229

0097-7403/11/$12.00 DOI: 10.1037/a0021384

Overt Attention and Predictiveness in Human Contingency Learning

M. E. Le Pelley
Cardiff University

Tom Beesley
University College London

Oren Griffiths
University of New South Wales

Two experiments used eye-tracking procedures to investigate the relationship between attention and associative learning in human participants. These experiments found greater overt attention to cues experienced as predictive of the outcomes with which they were paired, than to cues experienced as nonpredictive. Moreover, this attentional bias persisted into a second training phase when all cues were equally predictive of the outcomes with which they were paired, and it was accompanied by a related bias in the rate of learning about these cues. These findings are consistent with the attentional model of associative learning proposed by Mackintosh (1975), but not with that proposed by Pearce and Hall (1980).

Keywords: associative learning, attention, associability, eye tracking

One of the primary aims of associative learning theory is to establish the factors that determine how much we learn about a given stimulus or event under a given set of circumstances. Phrased differently, the question becomes: under a given set of circumstances, why do we learn more about some stimuli than about others?

One source of biases in learning, first established in studies of nonhuman animals (see Hall, 1991; Le Pelley, 2004, for reviews), and more recently studied in humans (e.g., Beesley & Le Pelley, 2010; Bonardi, Graham, Hall, & Mitchell, 2005; Kruschke, 1996; Le Pelley, Beesley, & Suret, 2007; Le Pelley & McLaren, 2003; Le Pelley, Reimers et al., 2010; Le Pelley, Schmidt-Hansen, Harris, Lunter, & Morris, 2010) relates to the experienced predictiveness of stimuli: that is, the ability of a cue stimulus to predict the occurrence of events of significance (such as rewards or punishments). Specifically, experience of the predictiveness of a cue seems to influence the rate of future learning about that cue, termed its *associability*.

Such findings have often been taken as support for models of learning in which the amount of learning that a given cue undergoes on a trial is modulated by a variable associability parameter, α (e.g., Kruschke, 2001, 2003; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall, 1980). This α parameter is itself a learned value which changes with experience of the cue's predictiveness. For example, Mackintosh's model (1975; see also Kruschke, 2001, 2003) proposes that a cue that is the best available predictor of the outcome with which it is paired (i.e., the cue with the highest predictiveness) will maintain a high α, while the α of all other cues will decline. In contrast, Pearce and Hall (1980; hereafter "the Pearce-Hall model") suggested that cues that are followed by unpredicted outcomes will tend to maintain a higher α than those that are followed by well-predicted outcomes. The relationship between predictiveness and associability in animals seems to be rather complex (see Le Pelley, 2004; Le Pelley, Turnbull, Reimers, & Knipe, 2010), but in studies of human contingency learning the typical finding is that cues previously experienced as having high predictiveness are learned about more rapidly than less predictive cues. This pattern is consistent with the view taken by Mackintosh's model, but not with the approach suggested by Pearce-Hall.

As an example, we will consider the study of human contingency learning by Le Pelley and McLaren (2003), which forms the basis of the current research. Table 1 shows the design of this study. Letters A–D and V–Y in Table 1 refer to different cues and o1–o4 to different outcomes, for example, "AV–o1" indicates that cues A and V were presented together, and that the correct outcome prediction was outcome o1.

During Stage I, cues A–D were predictive of the outcomes with which they were paired on each trial (A and D were consistently paired with o1; B and C with o2). Cues V–Y were nonpredictive, being paired equally often with o1 and o2. Over several blocks of Stage I training, participants learned to predict the correct outcome on each trial.

In Stage II, compounds of a previously predictive and previously nonpredictive cue were paired with novel outcomes o3 and o4. The objective statistical contingency between previously predictive cues (A–D) and Stage II outcomes was identical to that between previously nonpredictive (V–Y) cues and those same outcomes. For example, both A and X were paired with o3 an

This article was published Online First February 14, 2011.

M. E. Le Pelley, School of Psychology, Cardiff University; Tom Beesley, Division of Psychology and Language Sciences, University College London; Oren Griffiths, School of Psychology, University of New South Wales.

We thank John Patrick and the Communication Research Centre in Cardiff for the use of the eye-tracking equipment.

Correspondence concerning this article should be addressed to Dr M. E. Le Pelley, School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff CF10 3AT, United Kingdom. E-mail: lepelleyme@cardiff.ac.uk

Table 1
Design of the Experiments

Stage I	Stage II	Test
AV–o1	AX–o3	AC
AW–o1	BY–o4	BD
BV–o2	CV–o3	VX
BW–o2	DW–o4	WY
CX–o2		
CY–o2		
DX–o1		
DY–o1		

Note. Letters A–D and V–Y refer to different cues; o1–o4 refer to outcomes. On test, ratings of the cue compounds shown were obtained with respect to outcomes o3 and o4. All participants experienced all types of trial listed under a given stage of training.

equal number of times, on exactly the same trials. Despite this objective equivalence, in a final test, participants in Le Pelley and McLaren's (2003) study rated compounds AC and BD as significantly more predictive of outcomes o3 and o4, respectively, than were compounds VX and WY. This implies that, during Stage II, participants had learned cue–outcome associations more rapidly for previously predictive cues than for previously nonpredictive cues. That is, the associability of predictive cues A–D was higher than that of nonpredictive cues V–Y (see also Le Pelley, Suret, & Beesley, 2009).

This result is clearly consistent with the Mackintosh model's view of the relationship between predictiveness and associability. However, it is inconsistent with the view taken by the Pearce-Hall model. In its original formulation, the Pearce-Hall model uses a summed error term in the calculation of associability; as a consequence, what is crucial for determining associability is not how surprising the outcome is given the presence of an individual cue, but rather how surprising the outcome is given the combination (compound) of all currently presented cues. In the design shown in Table 1, all Stage I compounds are equally predictive of their respective outcomes, and hence the outcome occurring on each Stage I trial is equally surprising. Consequently, the original Pearce-Hall model predicts that all cues will have equal associability throughout the experiment. In a recent development of the model, Pearce and Mackintosh (2010) have suggested that associability might instead be governed by the predictiveness of each individual cue. However, since this model retains the fundamental principle of the Pearce-Hall model—that cues which are poor predictors of outcomes retain a high associability—this modified model is also unable to account for the findings of Le Pelley and McLaren's experiment, since it predicts that cues that are individually nonpredictive (i.e., cues V–Y) will maintain a higher associability than cues that are individually predictive (cues A–D) during Stage I. Hence the model incorrectly anticipates faster learning about V–Y than about A–D during Stage II.

The results of experiments such as that of Le Pelley and McLaren (2003) support the suggestion that the associability of a cue can be influenced by the previously experienced predictiveness of that cue. Some associability-based models go further, however, in suggesting that the associability of a cue is determined by the attention paid to that cue. Kruschke (2001, 2003) in particular was explicit in stating that (i) the amount of attention paid to a cue will be influenced by the experienced predictiveness of that cue, and (ii) this learned attention will influence the subsequent rate of learning about the cue, such that strongly attended cues will have a higher associability (and therefore will be learned about more rapidly) than weakly attended cues. Despite labeling his model "A theory of attention," Mackintosh (1975) remained agnostic on whether associability should be identified with attention; he argued that the mechanism underlying the model "may justify characterizing the present set of ideas as a theory of attention, but since that term has a number of connotations, it might be better to stress that what I am proposing is a theory about the associability of stimuli with reinforcement" (p. 294).

It is this idea that the term *attention* has a number of connotations that is of central importance. This is because paying greater attention to a cue would imply not only more rapid subsequent learning about that cue, but also all of the other consequences of attention previously established in the cognitive psychology literature (see Wright & Ward, 2008, for a review). The problem is that the vast majority of previous studies that have been taken as support for a relationship between attention and learning have relied entirely on rate of learning as a measure. As a result, this evidence cannot distinguish between an account in which predictiveness influences attention (as the term is understood by cognitive psychologists), and an account in which it merely determines the rate of learning about a cue (see Honey, Close, & Lin, 2010, for an example of this latter type of account, in which changes in learning rate are not driven by changes in attention).

The question, then, is whether changes in the associability of cues that result from differences in their predictiveness are accompanied by changes in the attention paid to those cues. In order to address this question, we need to measure the influence of predictiveness on aspects of stimulus processing (other than rate of learning) that have been established in the cognitive psychology literature as diagnostic of attention. Perhaps the most obvious feature of visual attention is that it tends to coincide with where our eyes are looking. It is, of course, possible to make *covert* shifts of attention that are not accompanied by eye movements (Posner, 1980). Nevertheless, eye movements and attentional shifts are generally tightly coupled (Deubel & Schneider, 1996), especially when dealing with relatively complex stimuli such as words.

Certain previous studies have examined the relationship between overt attention, as measured by eye gaze, and predictiveness. For example, studies of categorization by Rehder and Hoffman (2005a; 2005b) showed that participants spent more time looking at cues that were more diagnostic of category membership (that is, cues that were more predictive) than those that were less diagnostic. Related findings have been reported by Hogarth and his colleagues (Hogarth, Dickinson, & Duka, 2009; Hogarth, Dickinson, Hutton, Elbers, & Duka, 2006; Hogarth, Dickinson, Wright, Kouvaraki, & Duka, 2007); in studies of contingency learning using rewarding outcomes such as money, participants maintained attention for longer on cues that were more predictive of the delivery of this reward. Such findings of greater overt attention to more predictive cues are consistent with the view taken by Mackintosh's theory, but run counter to the Pearce-Hall model.

All of these previous studies measured differences in the overt attention to cues during training phases in which these cues differed objectively in their predictiveness. Recall from earlier, however, that Le Pelley and McLaren's (2003) study demonstrated a difference in the rate of learning about cues during a phase of the experiment (Stage II in Table 1) in which all cues were objectively equally predictive of the outcome with which they were paired; the only difference between these cues was in terms of their prior predictiveness during Stage I. In order to account for such findings, attentional theories must assume that an uneven attentional distribution arising as a result of differences in Stage I predictiveness can persist into Stage II, and hence bias what is learned away from the objective statistical cue–outcome relationships that are present in the data. These models therefore anticipate that differences in the associability of cues will be accompanied by differences in attention to those cues, but the studies of eye tracking described above do not allow us to assess this prediction.

Indeed, Hogarth, Dickinson, and Duka (2010) have recently argued that the attentional mechanism that determines the associability of a cue (which they term "looking-for-learning") might be quite separate from the attentional mechanism implicated in their studies using monetary rewards that are described above. Specifically, they argued that looking-for-learning will operate according to the principles of the Pearce-Hall model. This claim was based on the results of an experiment by Hogarth, Dickinson, Austin, Brown and Duka (2008), which used an evaluatively neutral outcome (a 50 dB tone). In their design, a compound of cues A and X was always followed by the outcome (AX+), C and X were never followed by the outcome (CX–), and B and X were followed by the outcome on 50% of presentations (BX+/–). Thus A and C were consistent predictors of reinforcement and nonreinforcement respectively, while B was a poor predictor. Differences in attention to these cues were assessed by comparing eye gaze to the unique cue on each trial (A, B, or C) with eye gaze to the common cue (X). Overall, Hogarth et al. (2008) found that participants spent longer looking at the unique cues than the common cue on all three types of compound trial. However, this bias in eye gaze was greater on BX trials, compared to AX and CX trials. The implication is that overt attention was greater to less predictive cues than to more predictive cues (since B was an inconsistent predictor of the outcome, while A was a consistent predictor of reinforcement and C was a consistent predictor of nonreinforcement).

The problem is that, once again, while Hogarth et al. (2008) showed a difference in overt attention to cues that differed in their objective predictiveness, they did not demonstrate that this difference in attention was related to a difference in the associability of the cues involved. That is, they did not assess the rate of novel learning about cues A, B, and C after the training described above, to verify that this too followed the predictions of the Pearce-Hall model. And indeed, as noted earlier, the majority of studies that have examined the influence of predictiveness on learning rate have instead found support for Mackintosh's theory (e.g., Bonardi et al., 2005; Kruschke, 1996; Le Pelley & McLaren, 2003).

To the best of our knowledge, only one previous study has measured the influence of predictiveness on both overt attention and associability within a single experiment. In a study of the blocking effect in contingency learning, Beesley and Le Pelley (2010) found that predictiveness did indeed exert a similar biasing influence on both eye gaze and rate of novel learning. Moreover, participants showing a larger influence of predictiveness on eye gaze also showed a larger influence of this variable on learning rate. However, we noted that these findings were explicable in terms of the relationship between predictiveness and attention suggested by both the Pearce-Hall and Mackintosh models, and hence could not decide between these different approaches.

Experiment 1

The aim of the current research was to examine the relationship between eye gaze and learning rate as measures of the influence of predictiveness, using a design that would allow us to distinguish between these different classes of attentional theory, and hence to test the claim made by Hogarth, Dickinson, & Duka (2010) that looking-for-learning operates according to the principles of the Pearce-Hall model. Specifically, these experiments used the "learned predictiveness" design of Le Pelley and McLaren (2003), shown in Table 1, while gaze location was monitored with an eye tracker. This allowed us to assess the general prediction of attentional theories of associative learning, that differences in the associability of cues during Stage II will be accompanied by differences in attention. More importantly, it enables us to establish whether any changes in overt attention that do occur follow the predictions of the Mackintosh model in particular, in which case we should observe greater overt attention to predictive cues.

Le Pelley and McLaren's original study used a rather concrete food allergy paradigm involving causal relationships between foods and allergies. Consequently, participants will be familiar with the cues and outcomes used in this design, and may have preconceived theories about causal relationships between them. In contrast, the previous studies of eye tracking cited above used more abstract procedures. For example, Hogarth et al. (2008) used arbitrary stimuli (snowflake-like visual patterns) with an auditory outcome (a white noise) in a predictive (rather than causal) learning design, making it unlikely that participants would bring any preconceived ideas to the experiment. In order to provide a closer comparison between our experiments and those reported previously, the current experiments used an abstract contingency learning procedure that was loosely based on these prior studies. Specifically, cues were nonsense words, and outcomes were distinctive sounds.

Method

Participants, apparatus and stimuli. Twenty-one Cardiff University students participated in exchange for course credit. The experiment was conducted using a Tobii 1750 Eye Tracker (Tobii Technology, Danderyd, Sweden)—a 17" (43.2 cm) monitor with a monitor-mounted eye tracker capable of recording eye gaze at a resolution of 50Hz. This offers a nonintrusive method of eye tracking and is able to compensate for small head movements. Participants sat approximately 60 cm from the screen. Stimulus presentation was controlled by a Visual Basic program, with timing determined by Windows API functions QueryPerformanceCounter and QueryPerformanceFrequency for millisecond resolution. Error signals were given over speakers.

The eight nonsense words, which acted as cues, were Conneastal, Dusapplity, Forditic, Holomoram, Luthinity, Miniputan, Pourplectly, and Slatorion. These were randomly assigned to letters A-D and V-Y in the design shown in Table 1 for each participant. The four outcomes were brief (≈.5s) sound clips of a splash, boing, squeak or zap. These were randomly assigned to outcomes o1–o4 for each participant.

Procedure and data analysis. Instructions stated that on each trial two words would be presented on-screen, that participants' task was to predict which sound would follow each pair of words, and that feedback would be provided. All responses during the experiment were to be made with the right-hand index finger.

Figure 1A shows a typical Stage I training trial. The two words were presented in white rectangles measuring 9.5 cm × 6.75 cm on screen (visual angle approximately 9° × 6.4°), which were arrayed vertically in the center of the screen, separated by 9.2 cm (≈8.8°). Participants made their prediction as to which of two sounds would occur using the left or right arrow keys to indicate o1 or o2, respectively. Once participants had made their response, a white frame appeared around the outcome they had chosen, and after a delay of 600 ms the appropriate outcome sound was presented over headphones. After a further 1000 ms the word "Correct" (in green) or "Wrong" (in red) appeared in the center of the screen as appropriate, and remained for 1300 ms, after which the trial ended and the screen was cleared. Cue words remained on-screen throughout the trial.

Following Stage I, participants were told that in the next training phase the same words would be presented in new pairings along with two new sounds, and that again their task was to predict which sound would occur after each pair of words. On each Stage II trial (see Figure 1B) the words were arranged horizontally in the

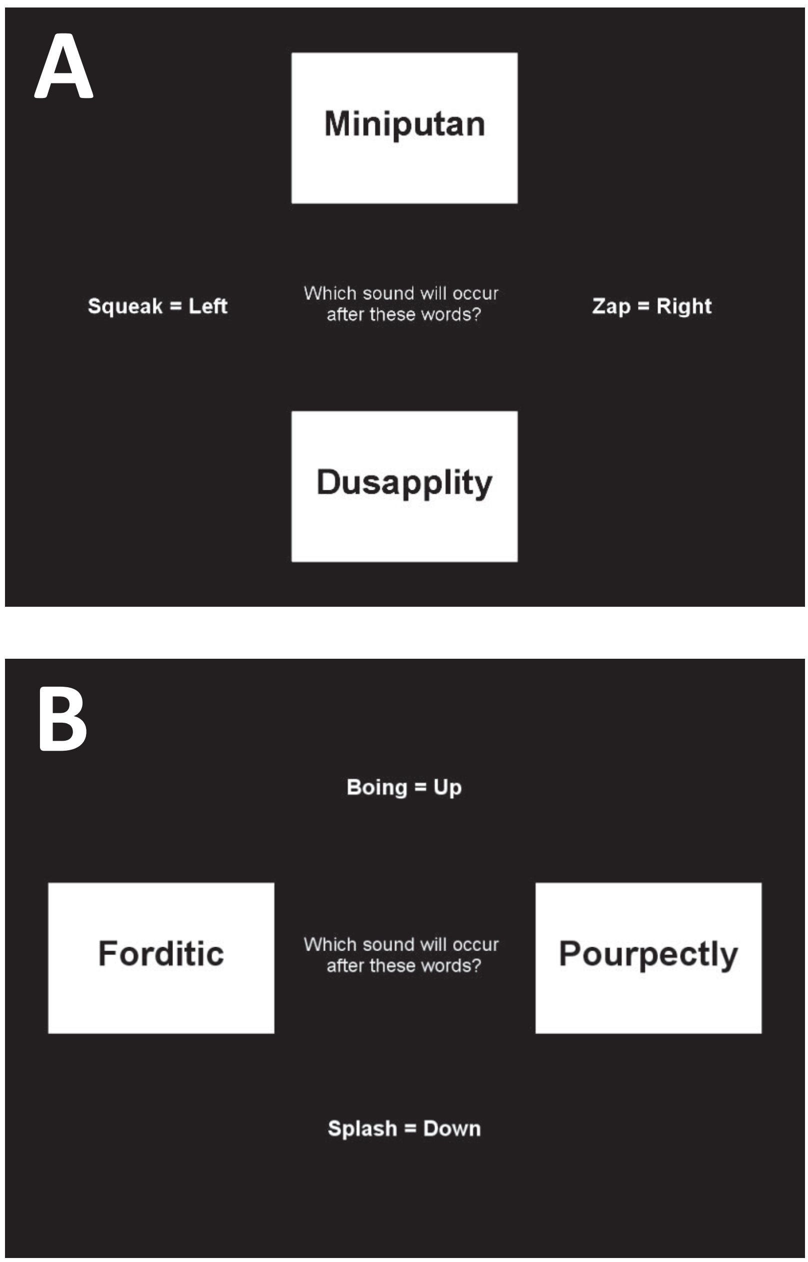

Figure 1. Screenshot of a typical training trial from (A) Stage I, and (B) Stage II of Experiment 1.

vertical center of the screen, separated by 11 cm (≈10.5°). Responses were now made with the up and down arrow keys to indicate o3 and o4, respectively. Combined with the fact that all responses in the experiment were made with the same finger, this ensures that responses in Stage II were orthogonal to those in Stage I. This was intended to reduce any noise resulting from carryover of knowledge regarding Stage I responses. That is, knowing that a particular cue was associated with a "left arrow" response in Stage I is of no use in deciding whether it will be associated with an up or down response in Stage II.

Stages I and II consisted of 18 and 6 blocks, respectively, with each of the trial types of each stage (see Table 1) occurring once per block. Trial order within a block was randomized, with the constraint that there could be no immediate repetitions across blocks. For each trial type, the left/right order of presentation of the cues in Stage I, and top/bottom order in Stage II, was counterbalanced across blocks. For example, for trial type AV–o1 in Stage I, there would be four presentations with cue A to the left of cue B, and four presentations with B to the left of A (the order of these presentations was randomized). The eye tracker was calibrated using a 9-point procedure at the start of the experiment and again after the 12th block of Stage I.

On the first test trial, two cue words appeared at the top of the screen, above a box containing the name of sound o3. Participants clicked on this box to play sound o3, which also brought up a rating scale from 0 (*very unlikely*) to 10 (*very likely*), which they used to rate how likely this sound was to follow the words shown at the top of the screen. On the immediately succeeding test trial, participants rated the same cues with respect to outcome o4. Participants provided ratings for all of the test compounds shown in Table 1 in similar fashion, and in a random order.

Data analysis. Following Le Pelley and McLaren (2003), these ratings were used to calculate difference scores for each compound. This was done by taking the rating for each compound with respect to the outcome (o3 or o4) with which its constituent cues were paired in Stage II, and subtracting from that the rating for the same compound with respect to the outcome with which its cues were not paired in Stage II. For example, the difference score for AC is given by the rating for AC with respect to o3 minus the rating for AC with respect to o4, because A and C were paired with outcome o3 during Stage II. Likewise, the difference score for BD is given by BD's rating for o4 minus its rating for o3, because B and D were paired with o4 during Stage II. These difference scores index the differential predictiveness of compounds with respect to Stage II outcomes—the extent to which a compound predicts the outcome with which it was paired in Stage II more than it predicts an outcome with which it was not paired. High difference scores (maximum = 10) indicate strong, selective performance, while a difference score of zero indicates no selective performance.

These difference scores are free from influences of generalization that would render any analysis based on raw rating data uninterpretable. Consider Stage I training, in which participants learn that A predicts o1, while X is not predictive of o1 or o2. Suppose that, immediately after Stage I, participants were asked how strongly cues A and X predict outcome o3. As o3 is novel at this point, neither cue will have a direct association with it. However, o3 has some similarity to o1 (both are sounds). Thus participants might generalize from the knowledge that A predicts o1 to also assume that it also predicts o3. Similarly, they might generalize from X's nonpredictive status with respect to o1/o2, to think that it will also not predict o3. So on the basis of generalization, and in the absence of further training, participants might perceive A as more predictive of o3 than is X. More generally, using raw ratings could generate a spurious difference in responding to previously predictive and previously nonpredictive cues that has nothing to do with differences in associability.

Using difference scores allows us to disentangle responding based on direct learning from that based on generalization. Suppose that, in our hypothetical single-stage experiment, we also asked people how strongly A predicted o4. Random assignment of sounds to outcomes o1–o4 in Table 1 ensures that, on average, o1 will be as similar to o3 as it is to o4, so that generalization from o1 to o3 is equal to that from o1 to o4. Therefore, if perception of A as a predictor of o3 were purely a consequence of generalization from its association with o1, we would expect an equally high rating for the A–o4 relationship (yielding a difference score of zero). If, however, participants rated A as a better predictor of o3 than of o4 (yielding a nonzero difference score), this would indicate that the A–o3 rating is not simply based on generalization, but that there is also a direct association between A and o3. The stronger this direct association, the greater the magnitude of the difference score.

Eye-gaze location was recorded every 20 ms for each eye independently. However, the eye tracker was occasionally unable to register gaze location for one or both eyes (e.g., as a result of blinks, head movements, etc.), resulting in missing gaze data. For each participant, the proportion of missing data for each eye was calculated across Stages 1 and 2, and gaze data from the eye with less missing data were used for all further analyses. Analyses of eye gaze reported below were based on dwell time, where the dwell time on a cue was defined as a recording of gaze within the white rectangular area surrounding the cue name (see Figure 1), and was summed across the period from cue onset to the detection of a valid response. Trials with response latencies greater than 10 seconds were not analyzed.

Results

The eye tracker could not be calibrated for two participants, so they were excluded from further analysis. We could only hope to observe an effect of Stage I predictiveness on learning and attention during Stage II if participants were able to learn the cue–outcome relationships during Stage I. Following Le Pelley and McLaren (2003), a selection criterion of 60% correct responses averaged across all Stage I trials was imposed (chance = 50% correct). One participant failed to achieve this criterion, and hence this participant's data were excluded from further analysis. For the remaining participants, the mean proportion of missing data for the "better" eye was 7.3% (SD = 3.2%).

Judgment data. Participants were clearly able to learn the correct cue–outcome relationships; mean percent correct across all trial types rose to 91.0% in the final block of Stage I training, and 91.7% in the final block of Stage II. The mean difference score—derived from test-phase ratings as described above—for compounds AC and BD (which both comprise cues that were predictive during Stage I) was 6.44, while that for compounds VX and WY (which both comprise cues that were nonpredictive during

Stage I) was 4.08. This difference was significant, $t(17) = 2.27$, $p < .05$.

Eye gaze data. Mean response latency for analyzed trials in Stage I was 2112 ms, and in Stage II was 2140 ms. Figure 2 shows mean dwell times on cues summed across the period from cue presentation to response. These data have been collapsed across all trials of each training stage, and are shown separately for predictive (A–D) and nonpredictive (V–Y) cues. Repeated measures analysis of variance (ANOVA) with factors of predictiveness and training stage revealed a significant main effect of predictiveness, $F(1, 17) = 14.4$, $p < .01$, with greater dwell time on predictive cues than on nonpredictive cues. The predictiveness $\times$ training stage interaction was nonsignificant, $F < 1$. Preplanned paired t tests revealed that dwell time on predictive cues was significantly greater than on nonpredictive cues during both Stage I, $t(17) = 3.07$, $p < .01$, and Stage II, $t(17) = 3.46$, $p < .01$. This confirms an attentional bias toward predictive cues in each stage. There was also a significant main effect of training stage, with greater dwell times in Stage I than in Stage II, $F(1, 17) = 37.8$, $p < .001$. This decrease in dwell time over the course of training presumably reflects participants' increased familiarity with the task in general, allowing them to make predictions more rapidly and thus decreasing the window over which dwell time is summed (notably mean dwell time also fell over the course of training within each stage, rather than just between stages). Moreover, increasing familiarity with the limited set of cues would allow participants to make these responses on the basis of an increasingly rapid assessment of the presented information; for example, if participants gradually learn that "Conneastal" is the only cue that begins with the letter C, they may move from reading the whole cue name on each trial to making a particular response as soon as they note that the cue name begins with the letter C.

It seems plausible that there will be individual differences in the extent to which participants distribute attention unequally between the elements of a stimulus compound (e.g., see Beesley & Le Pelley, 2010; Kruschke, Kappenman, & Hetrick, 2005; Le Pelley, Schmidt-Hansen et al., 2010; Wills, Lavric, Croft, & Hodgson, 2007). Recall that, according to attentional theories of learning, the influence of predictiveness on learning rate is mediated by attention. Specifically, those participants showing the greatest attentional advantage for predictive cues over nonpredictive cues during Stage II (assessed by eye gaze) should also show the greatest difference in learning about these cues (as assessed by rating difference scores).

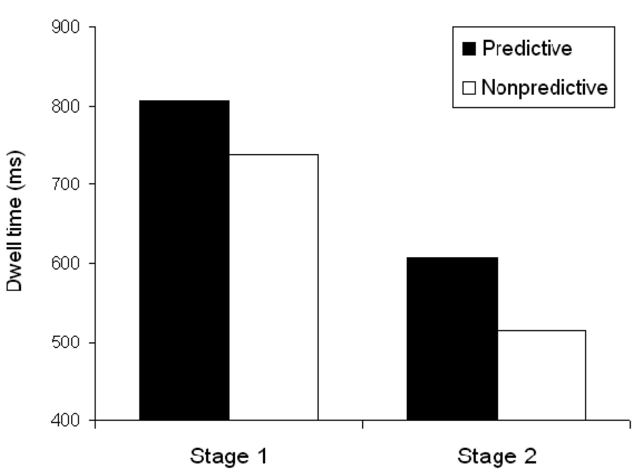

Figure 2. Preresponse dwell time on predictive and nonpredictive cues, averaged across the training blocks of each stage of Experiment 1.

For each participant, we calculated: (i) the predictiveness effect observed on our attentional measure (given by mean dwell time on predictive cues minus mean dwell time on nonpredictive cues, collapsed across all Stage II trials); and (ii) the predictiveness effect observed on our learning measure (mean difference score for compounds AC/BD minus mean difference score for compounds VX/WY). The correlation between these measures was significant, Spearman's $r_s(18) = .45$, $p = .032$ (one-tailed, since a positive correlation is explicitly anticipated).

Discussion

Consistent with the findings of Le Pelley and McLaren (2003), Experiment 1 revealed evidence for better learning about previously predictive cues than previously nonpredictive cues during Stage II. Thus, it would seem that the difference in the experienced predictiveness of these cues during Stage I induced a change in their associability, resulting in a biased pattern of learning.

Crucially, the eye-gaze data of Experiment 1 are consistent with the suggestion that this change in associability reflects a change in attention to the cues involved. The finding of greater overt attention to predictive cues than nonpredictive cues during Stage I (when these cues differed objectively in their predictiveness) agrees with the findings of prior studies (e.g., Rehder & Hoffman, 2005a) indicating that participants devote more attention to cues that are more diagnostic of category membership. The current data extend such findings by demonstrating that this unequal attentional distribution can persist into a training phase in which all cues have objectively equal predictiveness: a similar attentional bias toward previously predictive over previously nonpredictive cues was observed in Stage II, despite the statistical cue–outcome relationships being identical for both classes of cue. This pattern of results fits well with the predictions of attentional theories of learning, and provides a clear example of a statistically non-normative bias in overt attention. It was also found that those participants showing a greater bias in eye gaze during Stage II showed a correspondingly greater bias in what they learned about the cues involved, which further supports the general contention of attentional theories that biases in what is attended and what is learned are intimately related. This correlation is a pattern that we have replicated in unpublished work with the design used here, and in the context of the blocking effect (Beesley & Le Pelley, 2010).

The general finding of greater attention to, and more rapid learning about, predictive cues than nonpredictive cues follows the principles of Mackintosh's theory. While it is difficult to compare the sizes of effects at different points on the performance scale, Figure 2 indicates that the size of the attentional bias toward predictive cues in Stage II is at least as great as that in Stage I. This finding is again consistent with Mackintosh's theory, because it is the *learned* predictiveness of cues, rather than their objective predictiveness, that determines attention in this theory. Specifically, the model anticipates that the greater attention to predictive cues (e.g., cue A) following Stage I will fuel more rapid learning about these cues on the initial Stage II trials. Consequently, cue A will rapidly come to be perceived as more predictive of outcome o3 than is cue X. Since the Mackintosh model states that attention to cues with a higher learned predictiveness rises, while attention to cues that are perceived as poor predictors falls, the result will be that the attentional bias to cue A over cue X will further increase (or at least be maintained, if it has reached ceiling levels). Effectively, the Mackintosh model implements a positive feedback loop: a rise in attention produces faster learning, which in turn fuels a further rise in attention.

While the finding of greater attention to, and more rapid learning about, predictive cues than nonpredictive cues follows naturally from Mackintosh's theory, it is clearly inconsistent with the Pearce-Hall model. Consequently, our attentional data directly conflict with the suggestion made by Hogarth, Dickinson, & Duka (2010) that the looking-for-learning mechanism will follow the Pearce-Hall model. In some sense this is unsurprising. The general proposal of attentional theories of learning is that differences in associability will arise as a result of differences in attention and, to the best of our knowledge, all previous studies of associability effects in humans that are able to discriminate between these two models have favored Mackintosh's theory (see Le Pelley, Turnbull et al., 2010, for further discussion of this issue). Thus we might well expect, a priori, that attention, too, would follow the principles of this theory.

One question remains, however. Why did the relationship between predictiveness and attention in the current study (and those of Rehder & Hoffman, 2005a, 2005b) follow Mackintosh's theory, while in a seemingly similar study, Hogarth et al. (2008) instead found evidence consistent with the Pearce-Hall model? One possibility lies in a difference in the temporal arrangement of training trials. On each trial of Hogarth et al.'s experiments, a pair of cues was presented, along with an expectancy question asking participants to enter a numerical rating relating to how likely they thought an auditory outcome was to follow these cues. When participants had entered their rating, this expectancy question disappeared and the cues remained on-screen for a further 5 seconds. If the auditory outcome was scheduled to occur on this trial, it would then occur at a random point during the final 4 seconds of this period. Hence there was a variable and relatively long response–outcome delay, whereas in the current Experiment 1 (and Rehder & Hoffman's studies) this was fixed and short. Moreover, in Hogarth et al.'s studies, dwell time on cues was summed across a window of 5 seconds starting from cue presentation. This means that, if participants entered their expectancy rating within 5 seconds of cue onset, the dwell time window would also include a postresponse period (and could even include a postfeedback period). In contrast, in Experiment 1 (and Rehder & Hoffman's studies) dwell time was summed only across the window from cue onset to response. This raises the possibility that the attentional distribution between cues might change across the course of the trial, with a preresponse bias toward more predictive cues, and a postresponse bias toward less predictive cues. In Hogarth et al.'s data, the influence of the postresponse bias might outweigh the preresponse bias, producing an overall bias toward less predictive cues.

Experiment 2

These issues were investigated in Experiment 2, which used a variable response–outcome delay to allow us to investigate the attentional distribution across cues in pre- and postresponse windows. Given that the crucial focus of this experiment was the direct relation between predictiveness and attention, for simplicity Experiment 2 involved only the Stage I training phase from Table 1, which is most directly comparable to Hogarth et al.'s study (in that there are objective differences in the predictiveness of the cues under consideration).

Method

Participants, apparatus and stimuli. Fourteen Cardiff University students participated for course credit. Apparatus and stimuli were as for Experiment 1, and used a randomly chosen two of the four possible auditory outcomes from that experiment.

Procedure and data analysis. The procedure was as for Stage I training of Experiment 1, the only change being that the interval between participants' keypress response and presentation of the auditory outcome could be of duration 1200 ms, 2400 ms, 3600 ms, or 4800 ms. In each block of eight trials, each of these durations was used twice, in random order. Analysis of eye-gaze data was as for Experiment 1.

Results

For one participant, the proportion of missing data for the "better" eye was greater than 20% and so (following Beesley & Le Pelley, 2010) this participant was excluded from further analysis (this participant was also the only one failing to reach the criterion of 60% correct responding across Stage I). The mean percentage of missing data in the better eye for the remaining 13 participants was 11.3% ($SD = 4.4\%$). Stage I learning proceeded as expected, with mean percent correct rising to 94.2% in the final block.

Mean response latency for analyzed trials was 2353 ms. In a first analysis, dwell times on cues were summed across the period from cue presentation to response. These dwell times were then averaged across all training blocks, and across all response–outcome intervals (since prior to response all these trials are equivalent). Consistent with the results of Experiment 1, dwell time on predictive cues ($M = 775$ ms) was significantly greater than on nonpredictive cues ($M = 711$ ms) during this preresponse period, $t(12) = 2.71$, $p < .05$.

Figure 3 shows dwell times on cues for the period from response to the end of the trial (which includes the variable response–outcome interval, a subsequent fixed 1000 ms delay, and the 1300 ms during which feedback was displayed on-screen). Data are shown separately for each response–outcome interval, and are averaged across trial types and training blocks. Just as for the preresponse window, dwell time was greater to predictive cues than to nonpredictive cues. ANOVA with factors of interval (1200, 2400, 3600, and 4800 ms) and predictiveness (predictive cues vs. nonpredictive cues) revealed a significant main effect of predictiveness, $F(1, 12) = 10.6$, $p < .01$, that did not interact with interval, $F < 1$. Analysis of the related data summed only across the variable response–outcome interval (that is, the period following a response but prior to the presentation of any feedback)

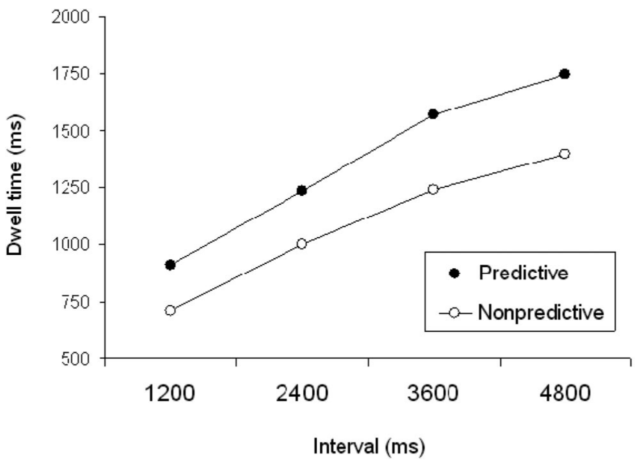

Figure 3. Dwell time on predictive and nonpredictive cues, averaged across the training blocks of Experiment 2. Data are shown separately for each response–outcome interval, and are calculated across the interval from response to the end of the trial (which includes the variable response–outcome interval, a subsequent fixed 1000 ms delay, and the 1300 ms during which the "Correct" or "Incorrect" feedback was displayed on-screen).

yielded a very similar pattern of results: a significant main effect of predictiveness, $F(1, 12) = 7.38$, $p < .05$ (with greater dwell time for predictive cues than nonpredictive cues), that does not interact with interval, $F(3, 36) = 2.29$, $p = .095$. In both of these analyses, the main effect of interval was also significant, smaller $F(1, 12) = 125.2$, $p < .001$, reflecting the fact that, for longer intervals, dwell time is summed over a longer temporal window.

Discussion

Experiment 2 replicated the finding of greater overt attention to predictive cues than to nonpredictive cues observed in the Stage I data of Experiment 1. Notably this attentional bias toward predictive cues persisted during the postresponse period. This would seem to rule out the possibility that the difference between the results of our experiments and those of Hogarth et al. (2008) merely reflect a difference in the window over which dwell time was summed. The data of Experiment 2 do not support the suggestion that the attentional distribution between cues changes across the course of the trial, with a preresponse bias toward more predictive cues, and a postresponse bias toward less predictive cues. Instead, there was an attentional bias toward more predictive cues regardless of the analysis window that was used.

General Discussion

Two experiments examined the relationship between predictiveness and overt attention in human contingency learning. Experiment 1 found an attentional advantage for predictive cues over nonpredictive cues during an initial training stage, and a corresponding difference in the rate of learning in a subsequent training stage during which all cues were (statistically) equally predictive of the outcomes with which they were paired. Notably, there was also a bias in overt attention toward previously predictive cues during this latter training stage, and there was a tendency for this attentional bias to be greater in those participants who showed a greater bias in learning.

These findings are clearly consistent with the dogma of attentional theories of learning; namely that predictiveness influences attention, which in turn influences rate of learning; in other words, that attention is a determinant of associability. More specifically, our data fit well with the relationship between predictiveness and attention suggested by Mackintosh (1975; see also Kruschke, 2001), wherein cues that are relatively good predictors of the outcomes with which they are paired tend to maintain greater attention than cues that are poorer predictors. In contrast, our data conflict with Pearce and Hall's (1980) theory, which anticipates equal attention to all cues in the current experiments, and Pearce and Mackintosh's (2010) recently proposed modification of this theory, which anticipates greater attention to nonpredictive cues.

The current results also conflict with the findings of Hogarth et al. (2008), who argued that their results implied an attentional bias toward less predictive cues, consistent with the Pearce-Hall model. Experiment 2 ruled out the suggestion that this discrepancy related to differences in the window over which dwell time was summed in these studies. So it remains unclear as to why such different results were found in the two cases. One possible answer relates to the complexity of the experimental designs. Our experiments were relatively complex, using eight different trial types in Stage I. Rehder and Hoffman's (2005a, 2005b) studies, which also found greater attention to more predictive cues, were more complex still, using many different trial types featuring cue compounds containing either three or four stimulus elements (rather than two in the current experiments). In contrast, Hogarth et al.'s study used only three different trial types (AX, BX, and CX), and hence placed considerably less cognitive load on participants. It has been suggested that differences in the load under which people operate might emphasize the role of different learning systems (De Houwer & Beckers, 2003; Dickinson, 2001; Le Pelley, Oakeshott, & McLaren, 2005). It would not seem unreasonable, then, to propose that cognitive load might also influence the type of attentional processes that modulate this learning. Under low load conditions, participants might aim to accurately establish the predictive status of every cue in the experiment, and hence devote maximal attention to those cues whose consequences are currently unclear (as suggested by the Pearce-Hall model). Under high load conditions, in contrast, participants may simply settle for getting as many correct answers as possible (and hence avoiding irritating error signals), and hence would pay attention to those cues which are most diagnostic of the correct answer on each trial (as predicted by the Mackintosh model). Future research will address this issue by manipulating the cognitive load imposed on participants in a learning task and observing any influence on the resulting pattern of attention as measured by eye gaze.

An alternative possibility is to ascribe the difference between the findings of these studies to the way in which eye gaze to predictive and nonpredictive cues was compared. In the present experiments (and those of Rehder & Hoffman) this comparison was direct, in that these cues were presented simultaneously on each trial. However, the approach used by Hogarth et al. to compare the attentional bias to, say, A (consistently reinforced) and B (partially reinforced) was indirect: it involved comparing attention to A versus X on AX trials, and B versus X on BX trials, and then comparing the magnitude of these two differences. It is possible that the pattern of eye gaze observed in this study was influenced by the presence of the common cue X on every trial, and that had A and B been presented simultaneously allowing for a direct comparison, a different pattern of attention may have been observed. In the absence of further evidence, however, this possibility remains speculation.

Attentional theories of learning make two key assumptions: (1) that predictiveness influences the attention paid to cues, and (2) that this attention in turn modulates the subsequent rate of learning about those cues (their associability). The current findings are certainly consistent with both of these assumptions. Indeed, the influence of objective differences in predictiveness on overt attention during Stage I demonstrates that the first assumption noted above must be correct. However, a note of caution is required with regard to the second of these assumptions. While our findings are consistent with the idea that attention modulates learning rate during Stage II, they do not provide unequivocal support for this conclusion. This is because Stage II of our experiment has a correlative design with regard to attention. In other words, we did not manipulate attention and look at the effect on rate of Stage II learning. Instead we manipulated predictiveness during Stage I, and observed the influence of this on the rate of Stage II learning. While we know that manipulating Stage I predictiveness has an effect on attention, we cannot be certain that it is this difference in attention that is driving the difference in learning during Stage II.

An example of an alternative account should make the problem clear. Suppose that (i) predictiveness during Stage I influences a nonattentional parameter, α, that determines the rate of learning about a cue (such that predictive cues maintain higher α than nonpredictive cues), and (ii) participants pay more attention to cues that they perceive as more predictive (as indicated by the Stage I data of the current experiments). The higher α of previously predictive cues A–D would promote more rapid learning about these cues during Stage II. This would cause these cues to be perceived as more predictive than are cues V–Y during Stage II, and consequently, according to the second of the above assumptions, participants would come to attend more to cues A–D than to cues V–Y. In contrast with the view taken by attentional theories of learning, however, this latter account supposes that it is a difference in learning rate that drives a difference in attention, rather than vice versa.

This problem of establishing the direction of causality is inherent in any attempt to test the hypothesized influence of attention on learning using a correlative technique. In order to be sure of a truly causal connection, an experimental approach is required. Suppose that there exists a manipulation that will reduce the extent to which a person is willing or able to use selective attention. If this manipulation also reduces the influence of experienced predictiveness on the rate of novel learning, this would support the suggestion that the influence of predictiveness on learning rate is via attention. Identification and testing of such manipulations remains a task for future research (see Le Pelley, 2010, for further discussion of this issue).

It is worth reiterating that, notwithstanding the above, the relationship between learned predictiveness and overt attention observed during Stage I of the current experiments does demonstrate a causal relationship. That is, changes in the learned predictiveness of cues caused changes in the overt attention paid to those cues, and these changes occurred in a manner that was consistent with the principles of Mackintosh's (1975) attentional theory of learning. Moreover, these changes in attention were accompanied by changes in learning rate that were also consistent with Mackintosh's model. Consequently, the current data support the suggestion that when associative learning theorists and other cognitive psychologists talk about "attention," they are indeed talking about the same thing.

References

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

Beesley, T., & Le Pelley, M. E. (2010). The impact of blocking on overt attention and associability in human learning. *Journal of Experimental Psychology: Animal Behavior Processes.* 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.

De Houwer, J., & Beckers, T. (2003). Secondary task difficulty modulates forward blocking in human contingency learning. *Quarterly Journal of Experimental Psychology, 56,* 345–357.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827–1837.

Dickinson, A. (2001). Causal learning: An associative analysis. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 54,* 3–25.

Hall, G. (1991). *Perceptual and associative learning.* Oxford: Oxford University Press.

Hogarth, L., Dickinson, A., Austin, A., Brown, C., & Duka, T. (2008). Attention and expectation in human predictive learning: The role of uncertainty. *Quarterly Journal of Experimental Psychology, 61,* 1658–1668.

Hogarth, L., Dickinson, A., & Duka, T. (2009). Detection versus sustained attention to drug cues have dissociable roles in mediating drug seeking behaviour. *Experimental and Clinical Psychopharmacology, 17,* 21–30.

Hogarth, L., Dickinson, A., & Duka, T. (2010). Selective attention to conditioned stimuli in human discrimination learning: Untangling the effect of outcome prediction, value and uncertainty. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 71–97). Oxford: Oxford University Press.

Hogarth, L., Dickinson, A., Hutton, S. B., Elbers, N., & Duka, T. (2006). Drug expectancy is necessary for stimulus control of human attention, instrumental drug-seeking behaviour and subjective pleasure. *Psychopharmacology, 185,* 495–504.

Hogarth, L., Dickinson, A., Wright, A., Kouvaraki, M., & Duka, T. (2007). The role of drug expectancy in the control of human drug seeking. *Journal of Experimental Psychology-Animal Behavior Processes, 33,* 484–496.

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: Oxford University Press.

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science, 8,* 225–247.

Kruschke, J. K. (2001). Towards a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45,* 812–863.

Kruschke, J. K. (2003). Attention in learning. *Current Directions in Psychological Science, 12,* 171–175.

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.

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

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.

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

Le Pelley, M. E., Oakeshott, S. M., & McLaren, I. P. L. (2005). Blocking and unblocking in human causal learning. *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 56–70.

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.

Le Pelley, M. E., Schmidt-Hansen, M., Harris, N. J., Lunter, C. M., & Morris, C. S. (2010). Disentangling the attentional deficit in schizophrenia: Pointers from schizotypy. *Psychiatry Research, 176,* 143–149.

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.

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.

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

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.

Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and a possible integration. In C. J. Mitchell & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 11–40) Oxford: Oxford University Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

Rehder, B., & Hoffman, A. B. (2005a). Eyetracking and selective attention in category learning. *Cognitive Psychology, 51,* 1–41.

Rehder, B., & Hoffman, A. B. (2005b). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 811–829.

Wills, A. J., Lavric, A., Croft, G. S., & Hodgson, T. L. (2007). Predictive learning, prediction errors, and attention: Evidence for event-related potentials and eye-tracking. *Journal of Cognitive Neuroscience, 19,* 843–854.

Wright, R. D., & Ward, L. M. (2008). *Orienting of attention.* New York: Oxford University Press.

Received April 21, 2010
Revision received August 3, 2010
Accepted August 9, 2010 ■