Journal of Experimental Psychology:
Animal Behavior Processes
2013, Vol. 39, No. 4, 323–333

0097-7403/13/$12.00 DOI: 10.1037/a0034073

The Fate of Redundant Cues in Human Predictive Learning

Metin Uengoer and Anja Lotz
Philipps-Universität Marburg

John M. Pearce
Cardiff University

In each of three experiments, a single group of participants received a sequence of trials involving pictures of a variety of foods presented individually or in pairs. Participants were required to predict in which trials the food would lead to a hypothetical allergic reaction. The different trials involved blocking, A+ AX+, and a simple discrimination, BY– CY+, in which each letter stands for a different food. Training trials were followed by a test in which participants were asked to predict how likely each kind of food would be followed by the allergic reaction. The principal purpose of the experiments was to determine how the redundant cue from blocking, X, would be judged relative to the redundant cue from the simple discrimination, Y. In contrast to predictions from currently influential theories of associative learning, X was regarded as a better predictor for the allergic reaction than Y.

Keywords: associative learning, redundant cues, blocking, simple discrimination

In a variety of experimental tasks, humans have been asked to indicate which of several concurrently presented cues is informative about the occurrence of a particular outcome. Such studies have frequently shown that participants can readily differentiate between informative and uninformative or redundant cues, and that rather little is learned about the significance of the latter. Two examples will serve to make this point.

Wasserman (1990) presented participants with a simple discrimination of the form AX+ BX–, in which they were told that a combination of two foods, say shrimp (X) and strawberries (A), would be followed by an allergic reaction, whereas another combination, say shrimp and peanuts (B), would be followed by no allergic reaction. In this task, A and B are informative by signaling the presence or absence of the allergic reaction, and X is redundant because it is less informative than A or B about the trial outcome. When subsequently asked how likely it is that X was responsible for the allergic reaction, participants indicated it was unlikely. A second group was trained with a pseudodiscrimination in which AX and BX were followed by an allergic reaction on half the trials and by no reaction on the remaining trials. X can be said to be at least as informative as A and B about the trial outcome in this task, and participants now stated it was possible that X was responsible for the reaction. In both cases, the relationship between X and the allergic reaction was the same, but by virtue of being redundant in the simple discrimination, it was afforded less significance than after the pseudodiscrimination. For a similar result with animals see Wagner, Logan, Haberlandt, and Price (1968).

The second example is provided by the phenomenon of blocking (e.g., Kamin, 1969). In an experiment by Aitken, Larkin, and Dickinson (2000). participants were first told that Food A is responsible for an allergic reaction, and then told that the combination of A and X together results in the same reaction. Here, A is informative, as a reaction occurs only when it is presented, and X is redundant because it provides no new information above that provided by A. Subsequent testing revealed that participants thought it was unlikely that X was responsible for the allergic reaction. Participants also received a control treatment in which the combination of two additional types of food, BY, signaled the allergic reaction. Test trials with Y then revealed that it was regarded as being more likely to result in an allergic reaction than X. The relationship between X and the allergic reaction and Y and the allergic reaction was the same, but making X redundant by virtue of the blocking treatment resulted in it gaining less influence than Y.

A common explanation for the foregoing pattern of results is that pairing a cue with an outcome results in the growth of an association between internal representations of these events. The strength of this association is then said to determine the perceived influence of the cue. Changes in the strength of the association are governed by a prediction error based on a summed error term. More precisely, the discrepancy between the combined associative strength of all the cues present on a trial, and an asymptotic value set by the magnitude of the outcome, is said to determine the change in associative strength of every stimulus on the trial in question. The first and perhaps most influential theory of this class was proposed by Rescorla and Wagner (1972). According to their theory, the increment in associative strength of Stimulus A, ΔV_A is given by Equation 1, in which α and β are learning rate parameters with values between 0 and 1, λ is the asymptote of conditioning determined by the magnitude of the outcome, and V_Σ

This article was published Online First September 2, 2013.

Metin Uengoer and Anja Lotz, Faculty of Psychology, Philipps Universität, Marburg, Germany; John M. Pearce, School of Psychology, Cardiff, United Kingdom.

The research reported in this article was supported by Grant UE 155/1-1 from the German Research Foundation (Deutsche Forschungsgemeinschaft) to Metin Uengoer.

Correspondence concerning this article should be addressed to Metin Uengoer, Faculty of Psychology Philipps-Universität Marburg Gutenbergstraße 18, 35032 Marburg Germany. E-mail: uengoer@staff.uni-marburg.de

is the sum of the associative strengths of all the cues that are present.

$$\Delta V_A = \alpha . \beta . (\lambda - V_\Sigma) \quad (1)$$

When Equation 1 is applied to blocking (A+ then AX+), provided that sufficient training is given with A+ to ensure it acquires asymptotic associative strength, then the introduction of AX+ trials will result in X acquiring no associative strength. On the other hand, if training is just given with BY+, then provided the two cues are of equal salience, they are each predicted to gain associative strength of .5λ. Thus, Equation 1 can readily account for blocking. As for the effects of the training with the simple discrimination, application of Equation 1 to an AX+ BX− discrimination leads to the predicted associative strength of A being .66λ, X being .33λ, and B being −.33λ. When Equation 1 is applied to the pseudodiscrimination, in which both AX and BX are paired with the outcome on half the trials, then the result predicted depends on the values assigned to the learning rate parameter, β. This value allows properties of the outcome to influence the rate of learning. If the value of this parameter is greater for trials with than without an outcome, then Equation 1 correctly predicts that the associative strength of X will be stronger after the pseudodiscrimination than after the simple discrimination.

An important, yet neglected, prediction from the foregoing analysis is that the redundant cue from a blocking treatment is predicted to gain less associative strength than its counterpart from a simple discrimination. On the basis of the calculations described, blocking is predicted to result in the redundant cue having no associative strength, whereas the cue that is present on both trials of the simple discrimination is predicted to have associative strength that is one third of the asymptotic value, λ. Moreover, this prediction holds true no matter what values are assigned to β on trials with and without an outcome.

At least informally, this prediction appears unlikely to be correct. It is true that blocking of the form A+ AX+ ensures that X is redundant, but at the same time, X is always paired with the outcome, and thus on trials when it is present, it can serve as an equally reliable, or informative, cue for signaling the outcome as A. In contrast, with a simple discrimination, CY− DY+, Y is not only redundant but it is also irrelevant, as it is followed by the outcome on some trials, but not others, in an unpredictable manner. Stimuli C and D, on the other hand, can be regarded as relevant because they can be used to predict trial outcomes accurately. It has long been appreciated that when presented together, relevant stimuli acquire more associative strength than irrelevant stimuli (e.g., Wagner, 1969), and for this reason alone, it might be expected that X, after the blocking treatment, will elicit a stronger response than Y, after the simple discrimination. If this prediction should turn out to be correct, then it would pose a serious challenge to the explanation offered by the Rescorla and Wagner (1972) theory for how humans attach more importance to relevant than irrelevant stimuli. Moreover, it is not just the Rescorla–Wagner theory that would be challenged by this finding. A number of other theories of associative learning assume that changes in associative strength are governed by a summed error term. They all make similar predictions as the Rescorla–Wagner theory about the effect of blocking and a simple discrimination on the associative strength acquired by the redundant cue (e.g., Gluck & Bower, 1988; Le Pelley, 2004; Pearce, 1994; Pearce & Hall, 1980).

To our knowledge, no study with humans has examined how the associative properties of redundant cues are affected by blocking and a simple discrimination. A single series of experiments has addressed this issue with rats and pigeons, and revealed that the redundant cue from blocking gains more associative strength than the redundant cue from a simple discrimination (Pearce, Dopson, Haselgrove, & Esber, 2012). If a similar effect can be found with humans, then, as noted previously, it will pose a challenge to a group of influential theories that have been used to understand learning in both humans and animals.

Experiment 1

The design of Experiment 1 is summarized in Table 1. In this experiment, we investigated whether, in human predictive learning, there is evidence that a redundant cue, X, from blocking, A+ AX+, is evaluated differently than a redundant cue Y from a simple discrimination, BY− CY+. We ran an experiment using a variant of the frequently used allergy task (e.g., Aitken, Larkin, & Dickinson, 2000). Participants were shown a sequence of pictures of different foods, either individually or in pairs, followed on each trial by information about whether they caused an allergic reaction. During the training trials, the participants had to predict what the allergic reaction would be and, during a final test stage, they had to indicate how likely it was that each of the foods was a cause of the allergic reaction. A within-subject design was employed, in which the trials from both tasks were intermixed—A+, AX+, BY–, CY+ (each letter represents a different food). If the effect described by Pearce et al. (2012) extends to humans, then during the test trials, X will be regarded as a more likely cause of the allergic reaction than Y.

In order to confirm that the A+ AX+ schedule was an effective means for producing blocking with X, and that the simple discrimination, BY– CY+, was an effective means for restricting learning about the significance of Y, additional trials were included in the experiment. Thus, the training stage contained trials with DE+, which were conducted in the same manner as for AX+. If intermixing A+ trials with AX+ trials results in blocking with X, then during the test stage, X will be rated as a less likely cause of the

Table 1
The Designs of the Experiments

Experiment	Stage 1	Test 1	Stage 2	Test 2
1	A+ AX+ BY– CY+ DE+ FG+/− HG+/−	A B C E F G X Y		
2	A+ AX+ BY– CY+	A B C X Y	A–B+	A B C X Y
3	A+ AX+ BY– CY+ D+/− DZ+/−	A B C D X Y Z		

allergic reaction than, say, E from the compound DE (e.g., Aitken et al., 2000). The training stage also contained a pseudodiscrimination, FG+/– HG+/–, in which G was paired with the allergic reaction according to the same intermittent schedule as Y from the BY– CY+ discrimination. If the simple discrimination in our experiment is effective at restricting the degree to which Y is rated as a cause of the allergic reaction, then during the test stage, the rating for Y will be lower than for G.

Method

Participants. Thirty-eight student volunteers (18 males) were tested individually and needed approximately 10 min to complete the experiment. Their mean age was 23 years. Participants received either course credits or a chocolate bar for participation.

Apparatus. Participants were tested individually in a room at the University of Marburg. Instructions and further necessary information were presented on a computer screen in German. The participants used a computer mouse for responding. Pictures of the following foods served as cues for the experiment: apples, bananas, cherries, grapes, lemon, carrots, pineapples, oranges, melons, and plums. For each participant, these foods were assigned randomly to the different letters A, B, C, D, E, F, G, H, X, and Y. The two outcomes were the occurrence of *stomachache* (+) and of *no stomachache* (–).

Procedure. Each participant was initially asked to read the following instructions on the screen:

> This study is concerned with the question of how people learn about relationships between different events. In the present case, you should learn whether the consumption of certain foods lead to stomachache or not.
>
> Imagine that you are a medical doctor. One of your patients often suffers from stomachache after meals. To discover the foods the patient reacts to, your patient eats specific foods and observes whether stomachache occurs or not. The results of these tests are shown to you on the screen one after the other. You will always be told what your patient has eaten. Sometimes, he has only consumed a single kind of food and on other times he has consumed two different foods. Please look at the foods carefully.
>
> Thereafter you will be asked to predict whether the patient suffers from stomachache. For this prediction, please click on the appropriate response button. After you have made your prediction, you will be informed whether your patient actually suffered from stomachache. Use this feedback to find out what causes the stomachache your patient is suffering from. Obviously, at first you will have to guess because you do not know anything about your patient. But eventually you will learn which foods lead to stomachache in this patient and you will be able to make correct predictions.
>
> For all of your answers, accuracy rather than speed is essential. Please do not take any notes during the experiment. If you have any more questions, please ask them now. If you do not have any questions, please start the experiment by clicking on the Next button.

The training stage consisted of eight blocks. Each of the seven different trial types (A, AX, BY, CY, DE, FG, and HG) was presented twice per block, yielding 112 trials. The order of presentation was determined randomly for each block. The two presentations per block of each of the trial types FG and HG were followed by the outcomes of stomachache on one occasion and no stomachache on the other.

Each trial started with the presentation of one or two food item pictures shown on a black background at the center of the screen. On trials with two food pictures, the stimuli were presented side by side, with the left–right allocation determined randomly on each trial. The sentences "The patient ate the following food(s)" and "Which reaction do you expect?" were presented above and below the cues, respectively. Participants made their predictions by clicking one of two response buttons shown side by side on the bottom half of the screen, The button on the left was labeled *no stomachache*, and the one on the right was labeled *stomachache*. Immediately after they responded, a feedback window appeared in the center of the screen. On no-stomachache trials, the statement "The patient has no stomachache" was presented as feedback, together with a green circle displaying a pictorial representation of a happy facial expression; on stomachache trails, the statement for the feedback was "The patient has stomachache," which was accompanied by a red circle displaying a pictorial representation of a sad facial expression. After clicking on the feedback window, the stimuli disappeared and the next trial started without further delay.

After completion of the training stage, participants read the following instructions:

> Now, your task is to judge the probability with which specific foods cause stomachache in your patient. For this purpose, single foods will be shown to you on the screen. In this part, you will receive no feedback about the actual reaction of the patient. Use all the information that you have collected up to that time.

On each test trial, one food-item picture was shown in the center of the screen, together with the question (presented above the cue), "What is the probability that the food causes stomachache?" Participants gave their ratings using a scale ranging from 0 (*certainly not*) to 10 (*very certain*). The rating scale was presented in the bottom half of the screen. The 11 values of the rating scale appeared side by side and participants chose one value by clicking on it. After participants confirmed their choice by clicking on an *OK* button presented below the rating scale, the next test trial started immediately. Participants did not receive any feedback during this stage. The eight cues (A, B, C, E, F, G, X, and Y) were each presented individually, twice, in a random sequence. For each cue, the two ratings were averaged for data analysis.

Results and Discussion

Figure 1 shows the percentage of participants who predicted the outcome of stomachache across successive presentations of each trial type throughout the training stage. During the course of training, the number of participants predicting stomachache increased for the trials in which this outcome occurred consistently, A+, AX+, CY+, and DE+, whereas the number decreased for trials that were never paired with this outcome, BY–. For the trials in which the outcome occurred intermittently, FG+/– and HG+/–, approximately half the participants predicted each trial would be followed by stomachache, whereas the remainder predicted no stomachache would occur. These intermediate rates of stomachache predictions resulted from an inconsistent pattern of responding within each participant. Collapsed across the 16 presentations of each of the trial types FG and HG, each participant predicted the

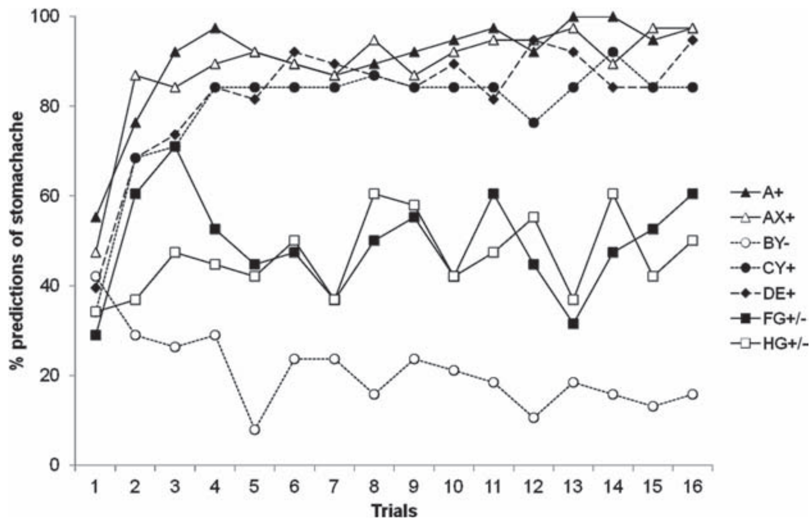

Figure 1. The group mean percentage of participants who predicted the outcome of stomachache after the seven trial types during the course of training in Experiment 1.

outcome of stomachache on approximately half of the trials (for FG, $M = 49\%$, $SD = 19\%$; for HG, $M = 46\%$, $SD = 19\%$).

The results from the test trials can be seen in Figure 2, which shows the group mean ratings of the likelihood that each cue (A, B, C, E, F, G, X and Y) would be followed by the outcome of stomachache. A one-way ANOVA of individual ratings for each of the eight stimuli revealed a significant difference among the stimuli, $F(7, 259) = 59.53$, $p < .001$, $MSE = 7.58$. As far as the informative cues from the blocking treatment and the simple discrimination are concerned, pairwise comparisons showed that the ratings for A were higher than for either C, $t(37) = 4.59$, or B, $t(37) = 17.12$, and that the ratings for C were higher than for B, $t(37) = 11.04$, $ps < .001$. Of most importance, the ratings for X were significantly higher than for Y, $t(37) = 6.79$, $p < .001$. Moreover, the ratings for E were significantly larger than for X, $t(37) = 2.09$, $p < .05$, and the ratings for G were higher than for Y, $t(37) = 5.22$, $p < .001$.

The results with the informative cues from the blocking treatment, A+ AX+, and the simple discrimination, BY– CY+, are in keeping with predictions from the theory by Rescorla and Wagner (1972). According to this theory, the repeated trials with A by itself serving as a cue for stomachache will result in it acquiring asymptotic associative strength and thus being regarded as a strong predictor of the outcome. This prediction is supported by the high ratings assigned to A during the test stage. Although C was consistently paired with the outcome, the presence of Y on every trial with C is predicted by the theory to overshadow C and to result in its associative strength being somewhat weaker than that

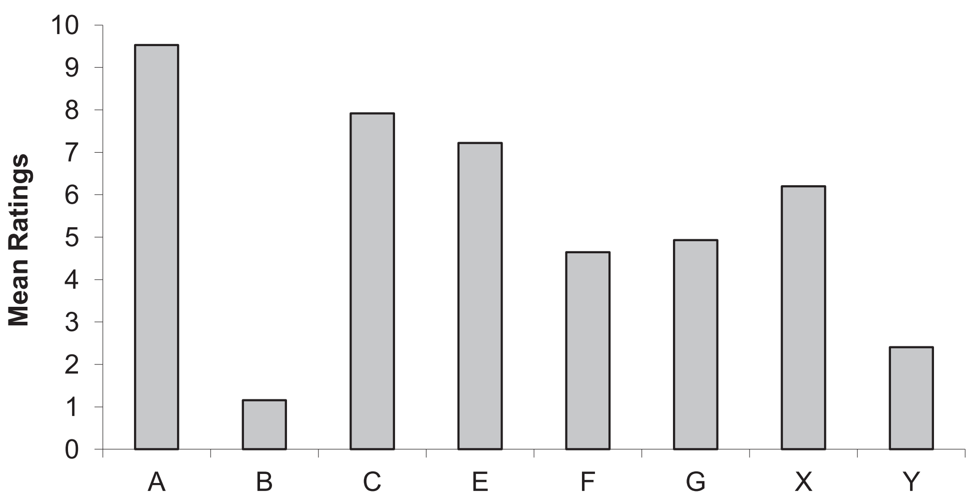

Figure 2. The group mean ratings of the likelihood that each of the eight foods would be followed by the outcome in the test trials of Experiment 1. Test trials were preceded by a training stage with A+ AX+ BY– CY+ DE+ FG+/– HG+/– trials.

of A. The results were consistent with this prediction. Finally, the theory predicts that B, by signaling the absence of the outcome in the presence of Y, will acquire negative associative strength, which would then explain its extremely low rating on the test trials. The results with the redundant cues are, in contrast, opposite of that predicted by the Rescorla–Wagner theory. Rather than being close to zero, the percentage of predictions for X was moderately high, and rather than being weaker than the rating for Y, the rating for X was significantly stronger than for this cue.

There were two additional, noteworthy findings from the experiment. One of these was the successful demonstration of the relative validity effect (e.g., Wagner, 1969), in which the common cue from the simple discrimination, Y, was given a considerably lower rating than the common cue, G, from the pseudodiscrimination, FG+/– HG+/–. The implication of this result is that the training with BY– CY+ was effective at restricting the rating given to Y. The second finding was that the redundant cue from the blocking treatment, X, was given a lower rating than one of the elements from the control compound DE+. The implication of this result is that the training with A+ AX+ was effective at restricting the rating given to X. Although both effects have been observed in other studies of predictive learning by humans (e.g., Aitken et al., 2000; Wasserman, 1990), they make clear for the present experiments that the ratings given to X and Y were not just determined by their own relationship with the outcome but also by the relationship of the stimuli that accompanied them with the outcome.

During the training stage, cues X and Y were not only paired with the outcome, they were also paired with other cues—X was paired with A, and Y was paired with both B and C. One explanation for the failure of Experiment 1 to reveal results consistent with predictions from the Rescorla and Wagner (1972) theory is that participants associated the redundant cues with the informative cues that accompanied them, and this affected the ratings given to X and Y. There is certainly evidence that such within-compound associations, as they have been called (Rescorla & Durlach, 1981), develop during training of the sort given in Experiment 1 (e.g., Dickinson & Burke, 1996; Melchers, Lachnit, & Shanks, 2004). Thus, on the test trial with X, participants might have recalled A, and as this cue was consistently paired with the outcome, it might have encouraged X to have been regarded as a reasonably strong predictor of the outcome. Conversely, the test with Y might have encouraged the recall of both B and C, and because one of these cues had been consistently paired with the absence of the outcome, participants may have treated Y, too, as being poorly associated with the outcome. The possibility is raised, therefore, that the strength of the associations between X and the outcome was weaker than between Y and the outcome, as the theory by Rescorla and Wagner (1972) predicts, but these differences were masked by the influence of within-compound associations. Experiment 2 was conducted in order to test this possibility.

Experiment 2

The two stages of training in Experiment 2 are summarized in Table 1. The first stage consisted of the same training and test trials as Experiment 1, A+ AX+ BY− CY+. The test confirmed the finding from Experiment 1 that the redundant cue from blocking, X, was regarded as being more likely to be followed by the outcome than the redundant cue from the simple discrimination, Y. In the second phase of the experiment, training was given with the intention of reversing the significance of A and B, by pairing B with the outcome of stomachache and A with the absence of this outcome. This treatment was then followed by the same test trials that were presented at the end of Stage 1. If the explanation for the results from the previous experiment in terms of within-compound associations is correct, then the results from the test trials at the end of Stage 2 will be opposite of that revealed by the test trials at the end of Stage 1. Thus, when X is presented for the test at the end of Stage 2, it may result in the recall of Cue A, but the recent treatment with this cue will mean it serves as a signal for the absence of the outcome, which will then weaken any tendency to regard X as being responsible for the outcome. In contrast, the presentation of Y will encourage the recall of both B and C, and because each of them can be regarded as a signal for the outcome, their recall may then enhance the degree to which Y is seen as being responsible for stomachache and result in it being regarded as a more accurate signal for this outcome than X.

Method

Participants. Twenty student volunteers (11 males) were tested individually and needed approximately 15 min to complete the experiment. Their mean age was 25.8 years. Participants received either course credits or a chocolate bar for participation.

Apparatus and procedure. The instructions were similar to Experiment 1. A subset of five food pictures from Experiment 1 served as cues for Experiment 2 (apples, bananas, cherries, grapes, and lemons) For each participant, these foods were assigned randomly to the different letters A, B, C, X, and Y.

Initial training consisted of 32 trials divided into eight blocks. Each of the four trial types (A, AX, BY, and CY) was presented once per block in a random order. This training stage was followed by a test stage similar to the one in Experiment 1, with each of the five cues (A, B, C, X, and Y) presented twice in a random order.

After the initial training stage and test, a second training stage followed. This second training stage was introduced by the instruction that the task is again to predict whether the patient will suffer from stomachache or not. During the second training stage, a reversal was implemented in which participants had to learn that cue A was not followed by stomachache (A–) and cue B was followed by stomachache (B+). This second training stage again consisted of eight blocks of trials with each block consisting of one presentation of each of the two trial types, A and B, shown in a random sequence. After the second training stage, a second test was given, which was identical to the first test. This second test was introduced by the instruction that the task is again to judge the probability with which specific foods cause stomachache in the patient.

Results and Discussion

The left half of Figure 3 shows the mean percentage of participants who predicted that stomachache would follow each trial type during the first stage of training. The results were similar to those obtained with the identical training in Experiment 1. The right half of the figure shows the results from the second stage of the experiment, and it is apparent that the reversal training with A and B was effective. By the end of this stage, stomachache was predicted by the majority of participants to be caused by stimulus B, but not by stimulus A.

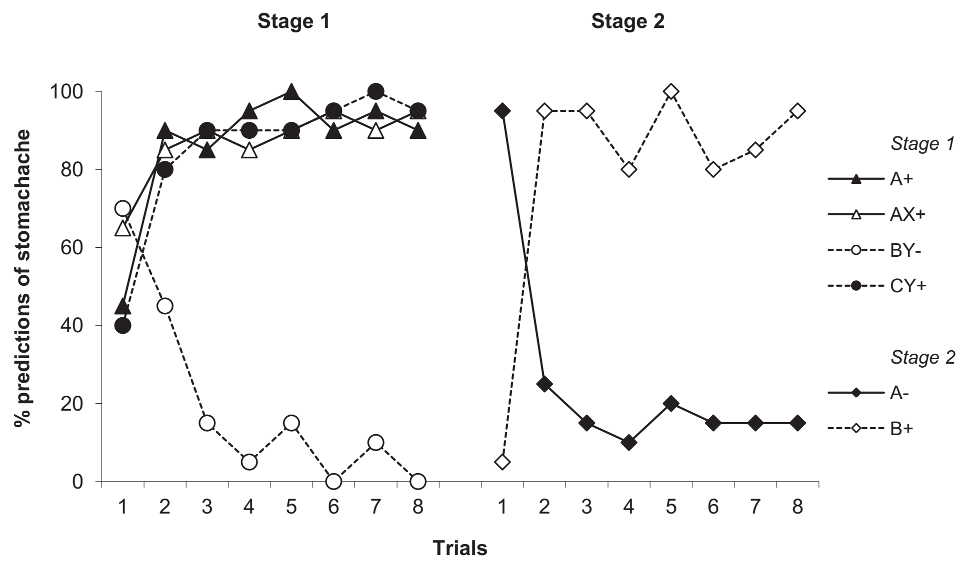

Figure 3. The group mean percentage of participants who predicted the outcome of stomachache after the four trial types of Stage 1, and the two trial types of Stage 2, in Experiment 2.

The results of the first and second set of test trials are portrayed in the upper and lower panels, respectively, of Figure 4. During the first test, ratings for A were considerably higher than for B, whereas the opposite relationship was observed during the second test. In support of this observation, a two-way ANOVA revealed a significant Stimulus (A and B) × Test Trial interaction, $F(1, 19) = 111.29$, $p < .001$, $MSE = 9.38$. Simple main effects of stimulus at each level of test trial showed that A was rated higher than B in the first test, $F(1, 19) = 580.56$, $p < .001$, and that B was rated higher than A in the second test, $F(1, 19) = 19.82$, $p < .001$. To return to the ANOVA, there was a significant effect of stimulus, $F(1, 19) = 7.82$, $p < .05$, $MSE = 7.39$, but not of test trial, $F < 1$.

Turning to the most important results, X was rated higher than Y in both tests. An ANOVA with the factors of stimulus (X and Y) and test trial showed a significant effect of stimulus, $F(1, 19) = 6.57$, $p < .05$, $MSE = 10.14$, but the effect of test trial, and the interaction were not significant, $F\text{s} < 1$.

The results from Stage 1 of Experiment 2 were the same as those from Experiment 1. Participants rated the redundant cue X from the A+ AX+ blocking treatment as being a more likely predictor of the outcome than the redundant cue Y from the a simple BY– CY+ discrimination. The novel conclusion to be drawn from the experiment is that the difference in the test ratings of X and Y does not appear to be determined by the associative strengths of A and B. Despite the successful reversal training with A and B in Stage 2, there was very little indication that it affected the response to X and Y. The implication of the findings from Experiment 2, therefore, is that the failure to confirm predictions from the Rescorla and Wagner (1972) theory in both experiments cannot be attributed to the influence of within-compound associations, as suggested in the discussion of Experiment 1.

The results from Experiment 2 also have implications for a further account of how within-compound associations influence performance, which is based on the extended comparator hypothesis (Denniston, Savastano, & Miller, 2001; Stout & Miller, 2007). At the heart of this hypothesis is the assumption that the associative strength of a stimulus is determined entirely by its relationship with the outcome that follows it. In the case of training with A+ AX+, therefore, providing sufficient trials are given, both A and X are predicted to gain asymptotic associative strength. The typical finding that X by itself elicits a weaker response than A is then attributed to the influence of a within-compound association between X and A. This additional association is deemed to be important because when X is subsequently presented alone, it will activate the representation of the original trial outcome by two routes: directly, through its own association with the outcome, and indirectly, through its association with A. The indirect activation is then said to counteract the direct influence of X and result in blocking being observed. Moreover, the greater the strength of the association between A and the outcome, the harder it will be for X to elicit a response. Thus, according to the extended comparator hypothesis, the reversal training with A in Stage 2 of Experiment 2 will weaken its association with stomachache and result in the rating for X being higher during the second than the first test. In contrast to this prediction, the experiment revealed the rating of X on the second test was no different than that for the first test, which implies that the extended comparator hypothesis is unlikely to provide a satisfactory explanation for our results. We now turn to consider an alternative explanation for the findings from the first two experiments.

Experiment 3

One possible reason for the higher ratings for X than Y in the previous experiments is that when X and Y were presented for testing, they retrieved memories of the training trials in which they were involved. Because trials with X were always followed by the outcome, and trials with Y were intermittently followed by the outcome, the retrieved memories might encourage participants to regard the outcome to be more likely to occur on tests with X by itself, rather than with Y by itself.

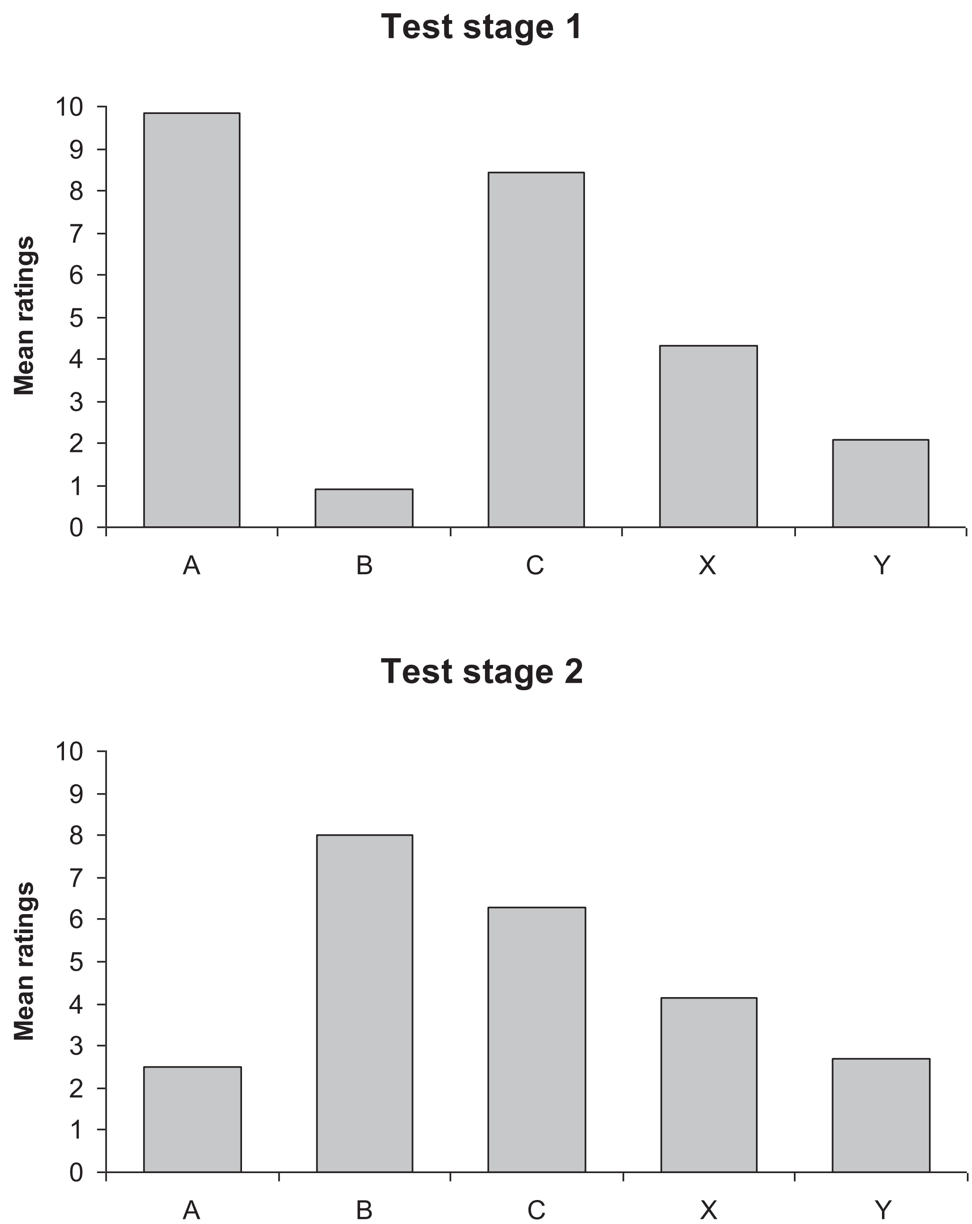

Figure 4. The group mean ratings of the likelihood that each of the five foods would be followed by the outcome in the two test stages of Experiment 2. Test Stage 1 was preceded by Training Stage 1 with A+ AX+ BY– CY+ trials. Test Stage 2 was preceded by Training Stage 2 with A– B+ trials.

Experiment 3 was conducted with this argument in mind. A single group received the same schedule of training as for the previous experiments, together with trials in which cue D was intermittently paired with the outcome either by itself, or in the presence of another cue, DZ, to create the schedule A+ AX+ BY– CY+ D+/– DZ+/–. Thus, like X, Z served as the redundant cue in a blocking treatment, and the question of interest is how participants would rate Z when it was presented for testing. On the one hand, because Z and Y were paired intermittently with the outcome during training, then, during testing, they may both retrieve memories of the presence and absence of stomachache and be rated as a poorer cause of this outcome than X. On the other hand, whenever Z was presented during training, it was just as informative about the trial outcome as the cue that accompanied it, D, whereas this was not the case for Y, and, on this basis, the rating for Z might be greater than for Y and similar to X. To test these contrasting predictions, the experiment concluded with test trials with all seven stimuli presented individually.

Method

Participants. Twenty student volunteers (five males) were tested individually and needed approximately 10 min to complete the experiment. Their mean age was 20 years. Participants received either course credits or a chocolate bar for participation.

Apparatus and procedure. The instructions were the same as those for Experiment 1. Pictures of apples, bananas, cherries, grapes, lemons, carrots, and pineapples served as cues for Experiment 3. For each participant, foods were assigned randomly to the different cues.

The training phase consisted of five blocks in each of which the six different trial types (A, AX, BY, CY, D, DZ) were presented twice in a random order. The two presentations per block of D, and of DZ, were followed by the outcomes of stomachache on one occasion and no stomachache on the other. After the fifth block, participants were shifted to the test phase. The test was conducted in a similar manner to the previous experiments, with each of the seven cues presented twice in a random order.

Results and Discussion

The percentage of participants who predicted stomachache would follow each of the six trial types, across the 10 trials of training, can be seen in Figure 5. By the end of this stage, nearly all the participants made correct predictions concerning the trials in which the outcome of stomachache occurred consistently, A+, AX+, and CY+, and for the trials in which this outcome never occurred, BY–. For the trials in which the outcome occurred intermittently, D+/– and DZ+/–, approximately half the participants predicted each trial would be followed by stomachache, whereas the remainder predicted no stomachache would occur. These intermediate rates of stomachache predictions resulted from an inconsistent pattern of responding by the participants. Collapsed across the 10 presentations of each of the Trial Types D and DZ, each participant predicted the outcome of stomachache on approximately half of the trials (for D, $M = 46\%$, $SD = 20\%$; for DZ, $M = 48\%$, $SD = 21\%$).

The results from the test trials with A, B, C, D, X, Y, and Z can be seen in Figure 6. The most important point to note is that X and Z were rated similarly and more highly than Y. A one-way ANOVA based on individual ratings for each of the seven cues revealed that different ratings were awarded to the single cues during the test, $F(6, 114) = 52.34$, $p < .001$, $MSE = 7.49$. Pairwise comparisons confirmed that the ratings for X were significantly higher than for Y, $t(19) = 3.28$, $p < .01$, and that ratings for Z were significantly higher than for Y, $t(19) = 3.33$, $p < .01$. The ratings for X and Z were not significantly different, $t(19) = .65$.

The test results with Cue Z demonstrate that a stimulus, which is made redundant by means of a blocking treatment, will be rated as a moderately strong cause of an outcome even when it is paired with the outcome on an intermittent schedule. The fact that the rating for Y was weaker than for Z, which was also paired with the outcome intermittently, suggests that the weak rating for Y was not just a consequence of its relationship with the outcome. Instead, the results point to the conclusion that one influence on the strength of the rating for a redundant cue is determined by its relative informativeness concerning the trial outcome. If the cue is less informative than the ones that accompany it, which was the case for Y, then it will be rated as a weak cause of the outcome. On the other hand, if the redundant cue is equally informative as the cues that accompany it, then it will be rated as a relatively strong cause of the outcome.

One puzzling finding from the experiments is that the trials with D+/– and DZ+/– resulted in a similar rating to D and Z when they were tested separately. The training with D+/– DZ+/– can be regarded as a blocking treatment, which should result in Z being regarded as of less significance than D in the same way that the A+ AX+ trials resulted in X being treated as less important than A. Of course, the blocking treatment with X involved a continuous reinforcement schedule, whereas with Z it involved partial reinforcement. It is possible that the use of an intermittent schedule reduced the rate at which training with D and Z progressed toward asymptote, in which case a stronger response to D than Z might have been observed if more training trials with these cues had been administered.

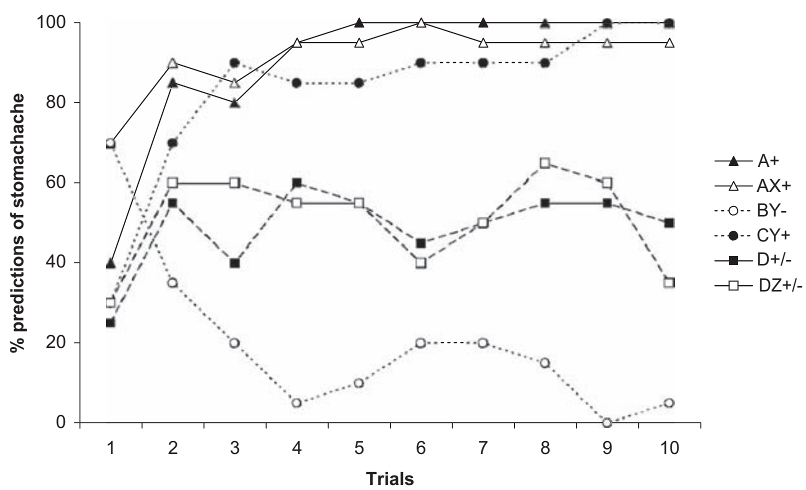

Figure 5. The group mean percentage of participants who predicted the outcome of stomachache after the six trial types during the course of training in Experiment 3.

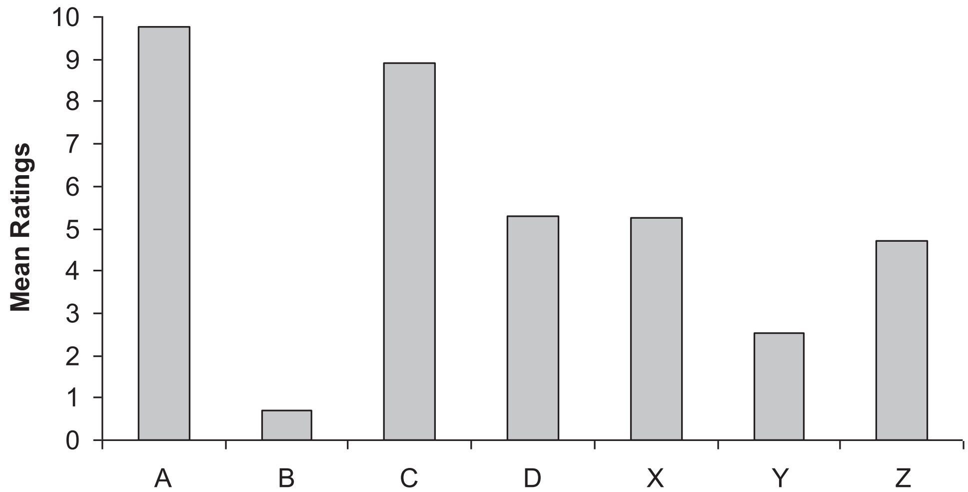

Figure 6. The group mean ratings of the likelihood that each of the seven foods would be followed by the outcome in the test trials of Experiment 3. Test trials were preceded by a training stage with A+ AX+ BY– CY+ D+/– DZ+/– trials.

General Discussion

In three experiments participants rated the redundant cue, X, from a blocking treatment, A+ AX+, as being of greater significance than the redundant cue, Y, from a simple discrimination, BY– CY+. When viewed informally, the results may not be surprising. The partial reinforcement schedule associated with Y, for the simple discrimination, means that, by itself, it is less informative about each trial outcome than the stimuli that accompany it, B and C. It would therefore make sense for participants to disregard Y and to focus on B and C. When blocking is concerned, the consistent pairing of X with the outcome ensures that on AX trials it is as informative as A about what will happen next, and participants may then be inclined to pay more heed to X than to Y. Although our results may be easy to understand in this informal way, it is not so easy to provide a formal explanation for them.

Much of the theoretical discussion concerning the experiments has been based on the influential theory by Rescorla and Wagner (1972). A superficial discussion of this theory in the introduction of this article revealed that it predicts the redundant cue from blocking will acquire less associative strength than the redundant cue from a simple discrimination. Before seeking alternative explanations for our results, we need to consider whether the initial interpretation of the theory is fully justified. Blocking experiments normally involve a period of training in which A, by itself, is paired with an outcome before the compound AX is paired with the same outcome (e.g., Aitken et al., 2000). Provided sufficient training is given with A, its associative strength is predicted to reach the maximum possible value, λ, and the associative strength of X will be zero. Given these values, the theory predicts that the redundant cue from blocking will gain less influence than its counterpart from a simple discrimination. In the present experiments, however, the trials with A+ and AX+ were intermixed from the outset of training, which makes it more difficult to draw a clear-cut prediction from the Rescorla–Wagner theory concerning our results. The parameter β in Equation 1 reflects the properties of the outcome, and it has been suggested that on trials with an outcome this value, βe, may differ from that on trials without an outcome, βi. In order to explore the significance of using different values of βe and βi for predictions from the Rescorla–Wagner model concerning A+ AX+ BY– CY+ training, Pearce et al. (2012) conducted a series of computer simulations based on Equation 1. When βe was the same or greater than βI, then the model consistently predicted that a test trial with Y will result in a stronger response than with X, at any point in training. That is, the opposite of our results is predicted. But when βe is less than βi, then the model predicts that during the early stages of training, the strength of the response to X will be stronger than to Y, but this relationship reverses as learning reaches asymptote. The results from Experiment 1 are of particular relevance to this discussion. The experiment revealed that the rating for the test trial with Y from the BY– CY+ discrimination was considerably less than for G from the FG+/– HG+/– discrimination. For the Rescorla–Wagner theory to explain this outcome, it has to be assumed that βe was greater than βi (see Rescorla & Wagner, 1972, pp. 84–86). Once this assumption is granted, the theory is then compelled to predict that the associative strength of X will always be weaker than of Y with our method of training, no matter how many trials are conducted before testing takes place.

A number of theories are based on an error-correction principle that is similar, or the same, as the one embodied in Equation 1 (e.g., Esber & Haselgrove, 2011; George & Pearce, 2012; Gluck & Bower, 1988; Le Pelley, 2004; Moore & Stickney, 1985). By incorporating this principle, they are then forced to predict that training with A+ AX+ will lead to A possessing asymptotic associative strength, and X having very little which, as we have just seen, makes it hard for them to explain the results. It does not follow from all theories of associative learning, however, that blocking will be as complete as predicted by Equation 1. Pearce (1987, 1994, 2002) has proposed that A+ AX+ training will result in representations of A by itself and of the compound AX entering into separate excitatory associations. When X is then presented by itself, it will elicit a response through stimulus generalization from AX, which will be weaker than that observed to A and AX, but stronger than if X had never been paired with the outcome. The theory thus goes some way toward explaining our results by predicting that the blocking treatment will leave X with a measure of excitatory strength. This does not mean, unfortunately, that the theory of Pearce is able to explain the outcome of the above experiments. Pearce (1994) has provided a set of equations that can be used to predict the strength of the response to X. As the theory was originally formulated, it follows from these equations that training with BY– CY+ will result in a stronger response to Y than X. Something else is needed, therefore, if the theory of Pearce is to explain the reported results.

An obvious addition to the theory is the possibility that during their exposure to a BY– CY+ schedule, participants learn to ignore Y because it is a less reliable predictor of the trial outcome than B and C. Successful demonstrations of the intradimensional–extradimensional shift effect, in which humans solve a discrimination more readily when the outcome is signaled by stimuli that were relevant, rather than irrelevant, to the solution of a previous discrimination lend support to this claim (e.g., Uengoer & Lachnit, 2012). If a loss of attention to Y is regarded as equivalent to a reduction in its salience, then it follows from the theory of Pearce (1994) that the rating for Y will be reduced and could then be weaker than to X. Although the suggestion that irrelevant stimuli are paid little attention has a long history (e.g., Sutherland & Mackintosh, 1971), we are making a novel proposal concerning the circumstances that result in a loss of attention. According to most theories, training with A+ AX+ is expected to result in a loss of attention to X because its associative strength is weaker than that of A (e.g., Le Pelley, 2004; George & Pearce, 2012; Mackintosh, 1975). The consequent loss of salience, according to the theory of Pearce (1994), will then result in X being treated in much the same way as Y after BY– CY+ training. To avoid this prediction, therefore, we suggest that through their exposure to A+ AX+ participants appreciate that A and X bear the same relationship with the outcome on AX trials, and thus continue to attend to X, at least to some degree. In contrast, on BY– CY+ trials, they come to appreciate that Y is a poorer predictor of the outcome than B and C, and thus attention to Y is reduced substantially. The obvious challenge for these proposals is to identify a mechanism that is responsible for producing these changes in attention to a stimulus. In the absence of a more convincing explanation for our findings, the challenge would seem to be worth pursuing.

The demonstrations that the redundant cue from blocking was given a stronger rating than the redundant cue from a simple discrimination are similar to findings from conditioning experiments with rats and pigeons (Pearce et al., 2012). It thus appears that our results are representative of learning in a wide range of species. Furthermore, the present results go beyond those described by Pearce et al. in two important ways. The experiments with rats and pigeons did not include a demonstration of the relative validity effect. It was thus not possible to rule out, conclusively, an explanation for the results with these species in terms of the theory by Rescorla and Wagner (1972), by assuming that the value of βe was less than βi. As noted, this assumption is not acceptable for the present experiments, which means they provide a more telling challenge to an interpretation in terms of the Rescorla–Wagner theory than those reported by Pearce et al. A further novel finding in the present report is that even when the blocking treatment was of the form D+/– DZ+/–, test trials still revealed that the redundant cue, Z, was rated as a stronger cause of the outcome than Y from the BY– CY+ discrimination. This comparison is important because it rules out the possibility that intermittently pairing Y with the outcome was solely responsible for the low rating it received in each experiment. Even though the experiments by Pearce et al. did not include trials of the sort D+/– DZ+/–, the present results suggest that the outcome of their experiments was not due to the intermittent reinforcement of the redundant cue from the simple discrimination.

References

Aitken, M. R. F., Larkin, M. J. W., & Dickinson, A. (2000). Super-learning of causal judgements. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 53,* 59–81. doi:10.1080/027249900392995

Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65–117). Mahwah, NJ: Erlbaum.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 49,* 60–80.

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: A model of attention in associative learning. *Proceedings of the Royal Society B, 278,* 2553–2561.

George, D. N., & Pearce, J. M. (2012). A configural theory of attention and associative learning. *Learning & Behavior, 40,* 241–254.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117,* 227–247. doi:10.1037/0096-3445.117.3.227

Kamin, L. J. (1969). Selective association and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42–64). Halifax, Nova Scotia: Dalhousie University Press.

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. doi:10.1080/02724990344000141

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

Melchers, K. G., Lachnit, H., & Shanks, D. R. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 57,* 25–53. doi:10.1080/02724990344000042

Moore, J. W., & Stickney, K. J. (1985). Antiassociations: Conditioned inhibition in attentional-associative networks. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 209–232). Hillsdale, NJ: Erlbaum.

Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review, 94,* 61–73. doi:10.1037/0033-295X.94.1.61

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review, 101,* 587–607. doi:10.1037/0033-295X.101.4.587

Pearce, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning & Behavior, 30,* 73–95. doi:10.3758/BF03192911

Pearce, J. M., Dopson, J. C., Haselgrove, M., & Esber, G. R. (2012). The fate of redundant cues during blocking and a simple discrimination.

Journal of Experimental Psychology: Animal Behavior Processes, 38, 167–179. doi:10.1037/a0027662

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

Rescorla, R. A., & Durlach, P. J. (1981). Within-event learning in Pavlovian conditioning. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 81–111). Hillsdale, NJ: Erlbaum.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Stout, S. C., & Miller, R. R. (2007). Sometimes competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review, 114,* 759–783.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York, NY: Academic Press.

Uengoer, M., & Lachnit, H. (2012). Modulation of attention in discrimination learning: The roles of stimulus relevance and stimulus–outcome correlation. *Learning & Behavior, 40,* 117–127. doi:10.3758/s13420-011-0049-z

Wagner, A. R. (1969). Stimulus validity and stimulus selection in associative learning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 90–122). Halifax, Nova Scotia: Dalhousie University Press.

Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology, 76,* 171–180. doi:10.1037/h0025414

Wasserman, E. A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science, 1,* 298–302. doi:10.1111/j.1467-9280.1990.tb00221.x

Received December 12, 2012
Revision received June 4, 2013
Accepted June 6, 2013 ■