

Final revision 09 May 2012

A Configural Theory of Attention and Associative Learning

David N George and John M Pearce

Short title: Attention and configural learning

Address for correspondence:

John M Pearce
School of Psychology
Cardiff University
Cardiff
CF10 3AT
pearcejm@cardiff.ac.uk

Abstract

A formal account of the relationship between attention and associative learning is presented within the framework of a configural theory of discrimination learning. The account is based on a connectionist network in which the entire pattern of stimulation presented on a trial activates a configural unit which then enters into an association with the trial outcome. Attention is assumed to have two roles within this network. First, the salience of stimuli at the input to the network can be increased if they are relevant to the occurrence of reinforcement, and decreased if they are irrelevant. Second, the associability of configural units can increase on trials when an outcome is surprising and decrease when an outcome is not surprising.

The theory in the present article differs in a fundamental way from the other theories considered in this special issue. When a pattern of stimulation is repeatedly paired with a given outcome then, according to the other theories, each element of the pattern has the capacity to enter into an association with the outcome. In contrast, the present theory assumes that each conditioning trial will provide the opportunity for only a single association to form. The association will be between a representation of the pattern in its entirety and a representation of the outcome with which it is paired. This configural account of conditioning has been presented formally on a number of occasions (Pearce, 1987, 1994, 2002; Pearce, Esber, George, & Haselgrove, 2008), and Pearce (2002) describes its application to a range of phenomena including blocking, overshadowing, conditioned inhibition, the solution of simple and complex discriminations, and so forth. A shortcoming with the original formulation of configural theory is that it failed to account for changes in attention that take place during the course of conditioning. We assume a change in attention to a stimulus is reflected by its conditionability. If the conditionability of a stimulus is high then it is assumed to be paid more attention than when its conditionability is low. The overall purpose of the present article is to present a formal modification of configural theory which allows it to explain how such changes in conditionability take place.

There is, in fact, a growing body of evidence to suggest that the attention paid to a stimulus is influenced according to two different sets of principles (see Pearce & Mackintosh, 2010). An early experimental design which has been said to show the influence one kind of attention on conditioning was conducted by Mackintosh and Little (1969). Two groups of pigeons initially received two discriminations of the form AW+ BW-, AX+ BX- (see Table 1). Each compound consisted of coloured lines. A and B belonged to one dimension, say orientation, which was relevant, and W and X to the other dimension, say colour, which was irrelevant. Once the initial discriminations had been mastered, subjects were transferred to a

set of new discriminations involving novel values from the two dimensions. Animals receiving an intra-dimensional shift (IDS), where the significance of the two dimensions remained the same for both stages, acquired the second set of discriminations more readily than those receiving an extra-dimensional shift (EDS), where the significance of the dimensions was reversed for the second stage. The generally accepted explanation for the foregoing outcome is that the initial training resulted in pigeons paying more attention to the relevant than the irrelevant stimuli, and that this effect generalised to the new values from the two dimensions for the test discrimination (e.g. Sutherland & Mackintosh, 1971; Mackintosh, 1975).

An experiment by Hall and Pearce (1979) revealed evidence of a rather different attentional process. Rats first received Pavlovian conditioning in which a CS was paired with a small shock before it was paired with a larger shock in Stage 2 of the experiment. On the basis of the principles that have just been put forward, it would be expected that the initial training would result in considerable attention being paid to the CS, which would then facilitate learning when the CS was paired with the larger US. In fact, conditioning in Stage 2 progressed slowly, which prompted Pearce and Hall (1980) to propose a new theory of attention. Animals were assumed to pay attention to stimuli whose significance was uncertain, such as a CS at the outset of conditioning, as this would promote rapid learning when it was needed. In contrast, once conditioning with a CS had reached asymptote then attention to it was assumed to decline, as there was nothing more to be learned about the stimulus, and hence no need to attend to it.

The two kinds of attention that have just been described can be viewed as serving rather different purposes. The first allows animals to detect stimuli that signal events of significance in their environment, and presumably will guide animals efficiently towards reward and away from punishment. This kind of attention could be said to alter the effective

salience of a stimulus, and thereby influence the chances of it being detected in the future. By enhancing the salience of relevant stimuli, and lowering that of irrelevant stimuli, this kind of attention would also make it easier to solve discriminations when irrelevant stimuli are involved. To take account of this type of change in attention we shall introduce a parameter, α , the value of which will reflect the effective salience of a stimulus.

The second kind of attention allows the mechanisms of associative learning to concentrate on stimuli whose significance remains uncertain and, about which, further learning is necessary. This kind of attention might be said to alter the conditionability of a stimulus. However, rather than alter the conditionability of individual stimuli, we propose this type of attention influences the conditionability of individual units in a connectionist network which represent the entire pattern of stimulation on the current trial. We shall introduce a second parameter, σ , to refer to the conditionability of a configural unit. In keeping with the proposals of Pearce and Hall (1980), σ is assumed to increase on trials when the outcome is surprising and decrease on trials when the outcome is expected.

Given these different functions of attention, and given there is good evidence for each of them (see Pearce and Mackintosh, 2010, for a review), it should not be surprising to discover that we are not the first to propose that they should both be incorporated into a single theory (e.g. LePelley, 2004). Indeed, Pearce, George and Redhead (1998) presented an outline of how a configural theory of associative learning might accommodate the two kinds of attention that have just been described, but it was presented informally. In the present article, therefore, we start by providing an introduction to the configural theory of Pearce (1994). We then show how this theory can be adapted formally to permit more attention to be paid to relevant than irrelevant stimuli and, at the same time, take account of the principles advocated by Pearce and Hall (1980). Finally, we explore how well the modified theory can account for

the basic phenomena of conditioning, particularly those which might be said to reveal an influence of attention on associative learning

A Configural Network for Discrimination Learning

Figure 1 shows a simple, configural connectionist network which can solve an AX+ BX- discrimination (ignore the dashed and dotted lines for the moment). In this discrimination, compound AX signals food, and BX signals the absence of food. Stimuli A and B are relevant by virtue of being uniquely associated with the presence and absence of food, respectively, and X is irrelevant by virtue of being uninformative about the trial outcome. To the left of the network is a column of receptor units that are activated whenever stimuli A, B or X are presented. Once a receptor unit is activated it will excite an input unit to the network. If the compound AX is presented to the network, the input units for A and X will be connected to a configural unit in the next layer. Since AX is paired with food, an excitatory link will develop between the configural unit and an output unit on the right-hand side of the network. If B and X are presented together and followed by nothing then a new configural unit representing this combination will be connected to the relevant input units, and it will enter into a negative or inhibitory association with the output unit. Once these connections are fully formed, presenting AX to the network will excite strongly the AX configural unit which will excite strongly the output unit. In addition, the presence of X will activate the BX unit, but the absence of B will mean this activation is incomplete. Even so, the modest activation of the BX configural unit will excite to some extent the inhibitory link to the output unit and weaken slightly the response excited by AX. If A should be presented by itself, it will activate only the AX unit, albeit incompletely, and lead to a relatively strong CR. Presenting B alone will activate only the BX unit and have an inhibitory effect on the output unit, while

presenting X alone will excite each of the configural units incompletely and result in a modest excitatory response.

Figure 1 about here

According to Pearce (1994) the formal principles behind these proposals are as follows. The level of activation of a receptor unit is determined by the intensity of the stimulus that excites it. The level of activation of an input is determined by the signal it receives from the receptor unit, and by the number of other input units concurrently activated. Equation 1 shows how the level of activation of input unit i , u_i , is determined when more than one stimulus is present (where i ranges from 1 to the number of input units activated).

$$\sum u_i^2 = 1 \quad 1$$

One consequence of Equation 1 is that the effective salience of a stimulus will be greater when it is presented alone, than when it is accompanied by another stimulus. It is partly for this reason that the theory of Pearce (1994) is able to explain the phenomenon of external inhibition (Pavlov, 1927). For ease of discussion we shall assume that whenever a stimulus is presented it activates its receptor unit to the maximum level of 1. If this level of activity is transferred to the input units, then it follows from Equation 1 that the value of u_i for any activated input unit will be given by Equation 2, where n is the number of activated input units.

$$u_i = \frac{1}{\sqrt{n}} \quad 2$$

Presenting a pattern of stimulation to the network for the first time will result in the recruitment of a single configural unit. The unit will become connected to all the input units excited by the pattern, and the configural unit will fire at its maximum level of 1. If the same pattern should be presented again then the same array of input units will be activated and the configural unit will again be activated to its maximum level. If a different pattern is presented

to the network, only a fraction of the units connected to the original configural unit will be excited and it will be only partially activated. Whenever a pattern of stimulation should fail to excite any existing configural unit to the maximum value, then the pattern can be regarded as novel and a new configural unit will be recruited.

The strength of the connection between input unit i and the configural unit j , $w_{i,j}$, is the same as the level of activation of the input unit, u_i . Equation 3 shows how the activation of configural unit j , a_j , will be affected when any pattern of stimulation, familiar or novel, is presented to the network (where i ranges from 1 to the number of input units).

$$a_j = \sum w_{i,j} . u_i \quad 3a$$

Pearce (1994) proposed that Equation 3a should be replaced by Equation 3b, in order to allow predictions from the network to be consistent with predictions from an earlier formulation of the theory (Pearce, 1987).

$$a_j = \left(\sum w_{i,j} . u_i \right)^2 \quad 3b$$

Adopting Equation 3b rather than 3a has at least two consequences. First, presenting a pattern to the network will activate the configural unit for a different pattern to a smaller extent according to Equation 3b than 3a. Thus the modification has the effect of sharpening the generalization gradient from one pattern to another. Second, Pearce (1994) has shown that the right-hand side of Equation 3b is formally equivalent to the expression $N_c^2 / (N_x \times N_y)$, where N_c is the number of common elements shared by patterns X and Y, and N_x and N_y are the number of elements unique to each of the patterns. This expression has been used on more than one occasion to compute the similarity between two patterns (e.g. Atkinson & Estes, 1963). Thus the extent to which one pattern of stimulation will excite the configural unit for another pattern can be said, according to Equation 3b, to be determined by the similarity between them. Kinder and Lachnit (2003, see also Pearce, Esber, George, & Haselgrove, 2008) have suggested that the exponent 2 in Equation 3b may not be fixed, but

might vary according to the nature of the experimental stimuli. For all the reported simulations, the exponent of Equation 3b was set at 2.

Although only one configural unit will be activated maximally on any trial, other units might be activated to a lesser extent as the result of previous training with patterns similar to the one currently presented. The previous trials may have resulted in these configural units entering into associations, of strength E , with the output unit. Pearce (1994) proposed that the overall activation of the output unit to pattern k , V_k , is then given by Equation 4 (where j ranges from 1 to the number of configural units).

$$V_k = \sum a_j E_j \quad 4$$

The only configural unit whose associative strength may change on a particular trial is the one representing the current pattern of stimulation, which will be the only one that is fully activated. The change in associative strength of this unit is given by Equation 5. In this equation, λ is the asymptote of conditioning, which is normally set at 1, and β is a learning rate parameter with a value set between 0 and 1 that is determined by properties of the reinforcer.

$$\Delta E_k = \beta.(\lambda - V_k) \quad 5$$

A computer simulation based on the foregoing equations revealed that the AX+ BX-discrimination will be solved when the associative strength of the configural unit for AX is 1.066, and that for BX is -.266. For the simulation, the value of λ was set at the maximum permissible value of 1, and β was set at .2. The simulation also revealed that, when presented alone, the stimuli A, B, and X would activate the output unit to .53, -.133, and .4 respectively. The negative associative strength of B will then allow it to manifest all the properties of a conditioned inhibitor (Rescorla, 1969).

Changes in salience (α) of relevant and irrelevant stimuli

The configural network is thus capable of solving discriminations involving both relevant and irrelevant stimuli. As the network has been presented, however, it is not possible to distinguish between those stimuli that are relevant, which should receive attention, and those that are irrelevant, which should be ignored. In order to achieve this goal we propose that the network shown in Figure 1 contains feedback links that transfer a signal from the configural layer to the input layer. These links are represented by the dashed lines. The purpose of the links is to enable input units for stimuli that in the past have been relevant to a trial outcome to be activated to a greater extent by their respective receptor units, than those that have been irrelevant. In the case of an AX+ BX- discrimination, a reduction in the sensitivity of the input unit for X will reduce the extent to which, say, compound AX will excite the configural unit for BX, and thus reduce generalisation between the two compounds. The discrimination will therefore be easier to solve than when there are no changes to the sensitivity of the input units. **Indeed, it was for this reason that we suggest that changes in salience take place at the input layer.**

In order to increase the sensitivity of an input unit, configural units that correctly predict the trial outcome feed back a signal that increases the sensitivity of the input units to which they are connected. On the other hand, configural units that incorrectly predict the trial outcome feed back a signal that reduces the sensitivity of their input units. More formally, on any trial, a signal is returned from each configural unit to all the input units to which it is connected. The magnitude of the feedback from configural unit j to input unit i , $f_{j,i}$, is the product of the associative strength of the configural unit and its net input activation, and is given by Equation 6a.

$$f_{j,i} = E_j \cdot \sum (w_{i,j} u_i) \quad 6a$$

As well as a magnitude, the feedback signal also has a sign. This is determined by the concordance between the associative strength of the configural unit and the occurrence, or not, of a US on that trial. If the value of E_j is positive, then feedback is positive when the US occurs, but negative if it does not occur. Conversely, if E_j is negative, feedback is negative when the US occurs, but is positive if it does not. In this way, positive feedback indicates that the configural unit is contributing to the prediction of the correct outcome for that trial, whereas negative feedback indicates that it is predicting the incorrect outcome.

The net feedback to each input unit is obtained by first calculating the sum of the signed feedback signals from each configural unit to that input unit. In order that the feedback reflects the extent to which each input unit contributes to the prediction of the trial outcome, this sum is then multiplied by the activation of the input unit. The net feedback to input unit i , F_i , is given by Equation 7 (where j ranges from 1 to the number of configural units).

$$F_i = u_i \cdot \sum f_{j,i} \quad 7$$

Once the feedback to each input unit has been calculated, the effectiveness of the input units is updated. If a single input is active, its effectiveness, α , is updated according to Equation 8a, where θ is a rate parameter (set to 0.1 for all of the simulations reported here). A single cue that predicts the correct outcome will gain in effectiveness, whereas a cue that predicts the incorrect outcome will lose effectiveness.

$$\Delta\alpha_i = \theta \cdot F_i \quad 8a$$

When two or more input units are active, then the change in the effectiveness of any one of these will be dependent not only upon the net feedback to that input unit, but also the feedback to other input units. This is achieved by subtracting the arithmetic mean of the feedback to all input units, \bar{F} , from the feedback to each unit, as shown in Equation 8b. Hence, it should always be the case that the effectiveness of a unit that is the best predictor of a trial outcome will increase, the effectiveness of the worst predictor should decrease, and

when all stimuli contribute equally to the prediction of the outcome there should be no changes in the effectiveness of any of them.

$$\Delta\alpha_i = \theta.(F_i - \bar{F}) \quad 8b$$

To understand the influence of an input unit's effectiveness on the network, we must first consider a modification to Equation 2 first described by Pearce (1994). This modification, given in Equation 9, was made in order to accommodate stimuli of different intensities, by allowing the activation of receptor units to have any real value rather than being confined to just 0 or 1. The activation of a particular input unit, u_i , is a function of the activation of its receptor unit, P_i , and that of all other receptor units. Despite the difference between Equations 2 and 9, the latter still ensures that the combined level of activation of all the input units on any trial is in keeping with the relationship expressed in Equation 1.

$$u_i = \frac{P_i}{\sqrt{\sum P^2}} \quad 9$$

The effectiveness, α , of an input unit affects activity in the input layer in much the same way as activity in its receptor units does. That is, an increase in the effectiveness of an input unit will have the same effect as an increase in the activation of its receptor unit, and a decrease in the unit's effectiveness will have the same effect as a decrease in the activation of its receptor unit. This is implemented in the manner shown in Equation 10, by simply multiplying the activation of each receptor unit by the effectiveness of the corresponding input unit. When the activation of each receptor unit in a given pattern is the same, and the effectiveness of the corresponding input units are equal, this equation reduces to Equation 2.

$$u_i = \frac{P_i.\alpha_i}{\sqrt{\sum (P.\alpha)^2}} \quad 10$$

Changes in conditionability (σ) of configural units

In order for the network to take account of the proposals of Pearce and Hall (1980), a further layer of feedback links has been included (see the dotted lines in Figure 1). Because we wanted to capture the idea of Pearce and Hall that the surprisingness of the US on one trial influences the conditionability on the next trial of the stimuli with which it was paired, these links influence the ease with which associations between configural units and output units can develop. In short, on every trial that a configural unit is fully activated its associability for the subsequent trial is determined by the degree to which the presented outcome is surprising.

Whereas Pearce and Hall (1980) proposed that the associability of individual stimuli was determined by the surprisingness of the events that followed them, we are proposing that the associability of entire patterns of stimulation is determined in this manner. More formally, the links transmit a signal of magnitude $(\lambda - V_k)$, which is determined by events on the current trial, to the configural unit that is maximally activated. This signal then determines the associability of the configural unit for the next occasion on which it is maximally activated. Equation 11 shows how the value of σ on trial n is determined by the events on trial $n-1$ (see Pearce, Kaye, & Hall, 1982). Equation 12 shows how σ influences the acquisition of associative strength to configural unit K.

$$\sigma^n = \gamma \left| \lambda^{n-1} - V_K^{n-1} \right| + (1 - \gamma) \sigma^{n-1} \quad 11$$

$$\Delta E_k = \sigma \cdot \beta \cdot (\lambda - V_k) \quad 12$$

Hence if the outcome on a trial with Pattern K should be unexpected, that is the discrepancy $(\lambda - V_k)$ is large, the conditionability of configural unit K will be high for the subsequent trial, but if the outcome is expected, that is $(\lambda - V_k)$ is small, the conditionability of this unit will be low. The parameter γ determines the extent to which σ is influenced by the discrepancy $(\lambda - V_k)$ on trial $n-1$, and the extent to which it is influenced by all previous trials. When γ is large, σ is primarily determined by the discrepancy on trial $n-1$.

Application of the network

In order to derive predictions from the network, a series of computer simulations was conducted. For these simulations the values of λ and β in Equation 5 were set at 1 and 0.1, respectively, and the values of the parameters θ and γ were set at 0.1. Unless stated otherwise, the intensity of a stimulus at the receptor units was 1, and its initial values of α and σ were both 0.6. The values of both α and σ were constrained within the range 0.05 – 1.0.

Conditioning with a single CS

Acquisition

The first simulation was conducted in order to determine the predicted course of acquisition when a single CS, A, is paired with a US of moderate intensity ($\lambda = .2$). This training will result in the recruitment of a single configural unit. Panel A of the top row of Figure 2 shows the predicted level of activity in the output unit, which will determine the strength of the CR, as training progresses. Panel B shows the predicted growth of the associative strength of the configural unit for A. Because only a single configural unit is involved in this training, the graphs in Panels A and B are identical. Panels C and D show, respectively, the predicted changes in the values of α and σ with continued training. As desired, the value of α increases as the A becomes a better predictor of the US, and σ decreases as the outcome of a trial becomes more accurately predicted.

Figure 2 about here

Hall-Pearce (1979) effect

The next simulation examined the predicted effect of pairing a CS with a large US, after it has been employed for the training just considered (Hall and Pearce, 1979). The

magnitude of λ for this new stage of training was 1. Panel A in Row 2 of Figure 2 shows that conditioning progresses slowly with A when it is first paired with the large US. This effect is a consequence of the initially low value of σ for the configural unit for A. Experience of the large US after A will, however, restore the value of σ and enable responding to attain a normal asymptote, as the figure shows. The results for a control CS, B, which was paired only with the strong US are also shown. The novelty of this stimulus at the outset of the simulation put it at an initial disadvantage in terms of its associative strength, but at an advantage through the high associability, σ , of its configural unit. The rapid conditioning with B then permits it to elicit stronger responding than during A after a few trials. Thus the modified theory is able to predict the finding of Hall and Pearce (1979) that pairing a tone with a weak shock results in latent inhibition when it is paired with a strong shock. Panel C of Row 2 shows that the value of α is predicted initially to be greater for CS A than CS B for this stage of the experiment. This difference, however, is not predicted to have an impact on the course of acquisition in the present design. If only a single stimulus is presented, then it follows from Equation 10 that no matter what its salience, as determined by either α or P , it will excite its input unit to its maximum value of 1. Changes in the rate of conditioning with this stimulus will then be determined solely by the value of σ . When two or more stimuli are presented together, then changes in α will be effective by influencing the degree of generalisation among compounds to which it belongs, and not by influencing directly the rate of association formation.

The latent inhibition effect just described is rather unusual because the pre-exposed stimulus was paired initially with a US. More conventional demonstrations of latent inhibition involve repeatedly presenting a stimulus by itself before it is paired with a US for the test phase (e.g. Lubow & Moore, 1959). The theory we are proposing predicts that latent inhibition will also be seen in these circumstances. When a stimulus is repeatedly paired with

nothing then it follows from Equation 11 that the value of σ for the configural unit it activates fully will decline and thus disrupt subsequent conditioning with that stimulus. It is worth noting that this treatment will not affect the value of α for the input unit activated by the pre-exposed stimulus. For a reduction in the value of this parameter to take place, the stimulus must be presented in a context where it is a less accurate predictor of the trial outcome than the stimuli that accompany it. In the case of conventional latent inhibition training, it can normally be assumed that all the stimuli on a pre-exposure trial signal the same outcome – nothing.

By acknowledging that pre-exposure takes place in an experimental context, the theory makes the additional prediction that latent inhibition will be context specific. Presenting stimulus A in context X, will result in a configural unit, AX, being connected to both stimuli. If this combination of stimuli is not followed by a US, then the value of σ for the configural unit representing AX will decline. Should A then be presented in a different context, Y, a new configural unit, AY, will be deployed. As the initial value of σ for the AY unit will be relatively high, conditioning with A in this context will progress more rapidly than in context X. A variety of experiments have confirmed this prediction (e.g. Hall & Channell, 1985). The latent inhibition effect demonstrated by Hall and Pearce (1979) is also predicted to be context specific; support for this prediction can be found in Swartzentruber and Bouton (1986).

Influence of partial reinforcement

The Pearce-Hall (1980) theory predicts that more attention will be paid to a stimulus that is paired with a reinforcer on a continuous rather than an intermittent schedule. To test this prediction, Kaye and Pearce (1984) presented a continuous group of rats with the sequence light-tone-food. A partial group received the same training but it was intermixed

randomly among trials with the light by itself. Subsequently, the light was paired with food and it was anticipated that appetitive conditioning would progress more rapidly in group partial than group continuous. Serial conditioning was used in Stage 1 for two reasons. The first was to reduce the level of magazine activity during the light to such a low level that a difference between the groups could be detected in Stage 2. The second was to ensure the light was followed by a stimulus of sufficient motivational significance to ensure the value of σ was greater in the partial group than the continuous group. These steps were effective as conditioning in Stage 2 progressed more readily with the partial than the continuous group (see also Swan & Pearce, 1988; Wilson, Boumphrey & Pearce, 1992).

The lower two rows of Figure 2 show the outcome of a simulation conducted in order to determine the predictions made by the present theory concerning the experiment of Kaye and Pearce (1984). The predicted effect of serial conditioning for both groups is shown in Row 3 of Figure 2. Panel B depicts the growth of the association between the configural unit for the light and the reinforcing properties of the tone. Panel A has been omitted from the row as it is hard to predict how the strength of the light-tone association will influence behaviour. In the study by Kaye and Pearce (1984) the conditioned response of magazine activity was weak in both groups during serial conditioning with the light. Panel C of Row 3 shows that the value of α is predicted to grow in both groups, but more rapidly for the continuous group. In keeping with theoretical predictions, the value of σ is predicted to be sustained at a higher value in the partial than the continuous group. The outcome of the simulation concerning the second stage of the experiment is presented in the bottom row of Figure 2, from which it is apparent that conditioning is predicted to be more effective with the partial than the continuous group. The transient increase in the value of σ in both groups (see Panel D) can be attributed to the fact that the delivery of the US immediately after the CS will initially be surprising.

Compound conditioning

Overshadowing

The next two simulations were conducted in order to explore the effect of using two stimuli presented simultaneously, AB, to signal a US. For the first of these simulations, the intensity of A and B was 0.6 and, for the second, A was set at the maximum value of 1 and the intensity of B was set at 0.5. Once again, only a single configural unit will be recruited, which will be activated fully whenever AB is presented to the input layer. Panels A and B of the upper row of Figure 3 show the familiar predicted pattern of a negatively accelerated growth in the strength of the response to the compound, and in the associative strength of the AB configural unit when A and B are of equal salience. The equivalent results for the simulation when A and B were of different salience can be seen in the next row. Panel C in both rows shows the predicted effect of compound conditioning on α . The value of this parameter is predicted to be maintained at its starting value for each stimulus when they are of equal salience, but when they are of different salience, α for the more salient stimulus is predicted to approach 1, while for the less salient stimulus it is predicted to approach zero. Panels D in both rows indicate that this training is predicted to result in σ increasing for a few trials and then declining to a low value for the configural units for the compound stimuli in both conditions. In order to determine the predicted effects of compound conditioning on the strength of the response elicited by individual stimuli, test trials were conducted at the end of the simulation. The associative strengths of both components of a compound were predicted to be .46 when their salience was .6. When the salience of A was 1 and B was .5, then the predicted response strength was .86 for A, and .21 for B. Thus reciprocal overshadowing is predicted for stimuli of equal salience. When they are of different salience, the strong stimulus overshadows considerably the weak stimulus, but the weak stimulus has little impact

on the strong one. These predictions are in keeping with reported findings (e.g. Mackintosh, 1976).

Figure 3 about here

Blocking

In order to investigate the predictions the theory makes concerning blocking, we examined the effects of pairing compound AB with a US ($\lambda = 1$), after extended training in which A by itself had been paired with the same US. Panel A of the third row of Figure 3 shows the predicted strength of the CR during the conditioning trials with AB. Responding is initially predicted to be weaker during AB than the asymptotic strength that would be expected for A. Such an outcome is a consequence of generalisation decrement that will result from presenting A in compound with B. Panel B shows that the associative strength of the configural unit for A is predicted to be sustained at a high level throughout conditioning with AB, while the associative strength of the configural unit for AB is predicted to be relatively weak. Panel C shows the effect of this training on α for the two stimuli. Its value remains consistently high for A, but is reduced to a low level for B. As for σ , its value is consistently low for the configural unit for A throughout the AB trials, which is understandable, but it is maintained at a moderate level to the unit for AB. The last result was not expected, but it is not hard to understand. As α for B declines rapidly during compound conditioning, the salience of this stimulus will eventually be so low that when AB is presented it will activate maximally the configural unit for A, rather than AB. As it is only possible for the value of σ for the maximally activated configural unit to change on any trial, from this point on there will be no further loss in the value of σ for the AB configural unit. A test trial with B by itself revealed that it is predicted to elicit a very weak response of strength .02. If the pre-training with A had been omitted, the above simulation of overshadowing with two stimuli of equal

strength indicates that the predicted strength of response to B would be substantially greater than for the blocking group, with a value of .46.

Figure 4 about here

Asymmetries in generalization

As it was originally formulated, the configural theory of Pearce (1994) predicted that a number of straightforward manipulations will have symmetrical effects. For example, a feature-positive discrimination, of the form AX+ X- is predicted to be acquired as readily as a feature negative discrimination, X+ AX-. Similarly, training with A+ and then testing with AX, is predicted to result in the same loss in the strength of responding as training with AX+ and testing with A. In fact, demonstrations of the feature-positive effect have shown that a feature-positive discrimination is acquired more readily than a feature-negative discrimination (e.g. Reberg & LeClerc, 1977), while Brandon, Vogel and Wagner (2000) have shown the decrement in responding brought about by the transition from AX+ to A is greater than that from A+ to AX. Both of these results can be explained with the modification to configural theory that is currently being proposed.

Feature-positive effect. Panel A of Figure 4 shows the predicted course of acquisition of a feature-negative, A+ AB-, and a feature-positive, CD+ C-, discrimination. Panel B shows the acquisition of associative strength predicted for the four configurations created by the two discriminations. Panel C shows the predicted changes in α paid to these stimuli as training progresses. Throughout training, α to A, B, and D is predicted to increase gradually, whereas it gradually declines to C. Such an outcome is understandable as the first three stimuli all provide information about the outcome of a trial, whereas this is not the case for C. As the theory predicts that the value of σ for the reinforced configurations will be the same throughout training the predictions concerning this parameter have been combined in Panel D; likewise for the nonreinforced configurations. The panel shows that, at least initially, σ for the

reinforced configurations is greater than for the nonreinforced configurations. This outcome is predicted because at the outset of training the discrepancy which determines the value of σ is greater on reinforced than nonreinforced trials. With continued training the value of σ declines to a low level for all four configurations.

Inspection of Panel A shows that the feature-negative discrimination is predicted to be acquired marginally more slowly than the feature-positive discrimination. The explanation for this difference resides in the loss of attention to C. This reduction in the effective salience of C will reduce generalization between the reinforced and nonreinforced trials, and facilitate the acquisition of the feature-positive discrimination. The absence of a similar loss of attention to the common element of the feature-negative discrimination, A, means that generalization between A and AB will be considerable throughout training and render this discrimination harder to solve than its feature-positive counterpart. Support for this analysis can be found in an experiment by Haselgrove, Esber, Pearce, and Jones (2010) in which rats received two feature-positive discriminations, AX+ X- and BY+ Y-, prior to an AY+ AX- BY- test discrimination. The AY+ AX- discrimination was acquired more slowly than the AY+ BY- discrimination, which implies that the salience of the previously relevant stimuli, A and B, was greater than of the previously irrelevant stimuli.

Figure 4 about here

Generalization between compounds and elements. We have already seen that conditioning with a compound, AB, with two stimuli whose salience is .6 will result in AB eliciting a CR of asymptotic magnitude, while either A or B alone will elicit a response of almost half this magnitude, .46. A further simulation was conducted to explore the predicted effect of the converse of this procedure. That is, of training with C by itself and testing with CD, where D is novel. If both stimuli are of equal salience (.6) and if training with C has

reached asymptote then the strength of the response to CD is predicted to be .70. The reason for this asymmetry in the effect of adding and removing a stimulus is based on the different changes in α that will take place during the initial training. The training with AB will result in each stimulus retaining its initial value of α of .6, whereas the training with C alone will allow α on this occasion to reach 1. In the case of AB, removing one component of a compound, which is equally salient to the other component, will result in a generalization decrement of approximately 50%. On the other hand, adding a novel stimulus with a value of α of .6 to a CS with a value of α of 1 will result in a generalization decrement of considerably less than 50%.

Asymmetries in acquisition during compound conditioning

We close this discussion of compound conditioning, with a result by Rescorla (2000) which, at first glance, appears to contradict our theoretical proposals. Animals received training of the form A+ C+ X+ XB- XD- before conditioning with AB+. Rescorla found that responding during subsequent test trials with BC was stronger than during AD, which led him to conclude that compound conditioning with AB resulted in an asymmetry in the rate of acquisition of associative strength by A and B, with the rate being faster for B than A. According to the theory of Pearce (1994), such an asymmetry should not have been found. The AB+ trials are predicted to result in the growth of a single association between a representation of the compound in its entirety and the reinforcer so that any apparent growth in the associative strength of A will be matched by a similar growth for B (see Rescorla, 2000, p. 436). Once it is acknowledged that the conditioning trials might influence the value of α for A and B, then the present version of the theory makes a different prediction. The value of α is predicted to be close to 1.0 for all five stimuli by the end of the initial training, but slightly lower to X than the other stimuli. The AB+ trials are then predicted to maintain this

value for A, and result in α falling for B because the AB configural unit will be paired with a different outcome to the other configural unit activated by B, XB. The loss of salience by B will then permit more generalisation of excitation, and less generalisation of inhibition, to BC than AD, and thus result in the outcome recorded by Rescorla.

Evidence for enhanced attention to relevant stimuli

There is a growing body of evidence in support of the proposal by Sutherland and Mackintosh (1972) that cues which are relevant to the solution of a discrimination are paid more attention than those that are irrelevant. The purpose of this section is to determine whether our modified version of configural theory is compatible with this evidence.

AX+ BX- Discrimination

We introduced Pearce's (1994) configural theory by demonstrating its application to an AX+ BX- discrimination. The predictions from the present revision of configural theory concerning this discrimination can be seen in Figure 5. The predicted changes in the strength of the response to the two compounds is shown in Panel A, and the predicted changes in the associative strengths of the compounds can be seen in Panel B. Panel C shows that the discrimination will result in a decline in the value of α for X, while the value of this parameter increases for both A and B. In support of these predictions, Dopson, Esber and Pearce (2010) and Pearce, Esber and Dopson (2008) have shown that after an AX+ BX- discrimination the associability of A is greater than of X, while Dopson, Williams, Esber and Pearce (2010) have shown that the associability of B is greater than of X. To our knowledge, this last report is the first to demonstrate that animals pay considerable attention to a cue that signals the absence of food.

Figure 5 about here

IDS-EDS Effect

We return now to the IDS-EDS experiment which was introduced at the outset of this article (see Table 1). During Stage 2 of this type of experiment, subjects are given discriminations with stimuli belonging to two dimensions. For the IDS group the dimension that is relevant in Stage 2 was also relevant for a discrimination with different values from the two dimensions in Stage 1. In contrast, for the EDS group the dimension that is relevant in Stage 2 was irrelevant in Stage 1. In order to derive predictions from the model concerning this type of design, a mechanism is required that permits greater generalization among stimuli that belong to the same dimension, but not among those belonging to different dimensions. It is not sufficient to assume that all the stimuli from the same dimension share a common element. Although this will permit generalization among these stimuli, it will also mean that during Stage 1 of an IDS-EDS experiment, the common element for the relevant dimension will be irrelevant. The consequent loss of attention to this cue will then prevent a beneficial effect of the Stage-1 training being observed in Stage 2, and the group trained with an IDS will perform similarly to the group trained with an EDS.

To circumvent the foregoing problem, we propose that stimuli can excite unique and common receptor units. A unique receptor unit will be excited by a specific stimulus, whereas a common receptor unit will be excited by two or more stimuli, even if they are presented individually. If it is assumed that two similar stimuli excite a fair number of the same common units, whereas only a few such units are excited by two very different stimuli, then it follows that generalization will be stronger between similar rather than different stimuli. It also follows that an increase in attention to one stimulus will generalize to similar stimuli. With these principles in mind, a computer simulation was conducted for the IDS-EDS design outlined in Table 1. For the purposes of simplifying the simulation, each stimulus was

assumed to excite one unique receptor unit and three common receptor units, each of which would be excited by just one of the three remaining stimuli belonging to the same dimension.

Panels A and B in the top row of Figure 6 show, respectively, the predicted changes in the strength of responding to the reinforced and nonreinforced compounds for Stage 1 of training, and the predicted changes in the associative strengths of the configural units for the four compounds. Panel C shows how α changes for the relevant and irrelevant input units during the initial discrimination training. As expected, the irrelevant units are predicted to lose attention and the relevant units to gain attention. Panel D shows the changes in the value of σ that are predicted for the configural units during the first stage of the experiment.

Figure 6 about here

The two panels in the lower row of Figure 6 show the predicted changes in the strength of responding during Stage 2 for the reinforced and nonreinforced compounds for the group trained with the IDS discrimination (left-hand panel), and the group trained with the EDS discrimination (right-hand panel). The former is predicted to be acquired more readily than the latter. George and Pearce (1999) have demonstrated successfully an IDS-EDS effect using a biconditional discrimination. Although the design of the experiment was substantially more complicated than for the study just considered, computer simulations again confirmed that the present theory is able to predict correctly this result.

Configural Discriminations

Many of the discriminations that have been studied can be solved by focussing on the significance of individual stimuli. For instance, a feature-positive discrimination, AB+ B-, can be solved once it is appreciated that A signals the outcome, and that B is of no significance. This type of discrimination can be referred to as an elemental discrimination. Other discriminations cannot be solved in this way. Instead, the solution depends upon the organism

grasping the significance of combinations of stimuli. For instance, a positive patterning discrimination, A- B- AB+ can be solved by learning that A and B separately signal the absence of the outcome, but together they signal the occurrence of the outcome. This type of problem is referred to as a configural discrimination. We now turn to the predictions our theory makes concerning configural discriminations. In fact, the distinction between elemental and configural discriminations is largely irrelevant for the theory, as all associations are assumed to be based on configural representations.

The upper row of Figure 7 shows the predictions made by the theory for the changes in associative strength for the reinforced and nonreinforced trials of a positive patterning discrimination. The discrimination is predicted in Panel A to progress quite normally. For this, and all of the remaining simulations, unless stated otherwise individual stimuli were assumed to excite unique, but not common receptor units. We should add that the addition of common receptor units to the simulations does not affect qualitatively the predictions that are made.

Figure 7 about here

The second row of Figure 7 shows the predicted outcome of a negative patterning discrimination, A+ B+ AB-. The course of acquisition of the discrimination, which is shown in the left-hand panel, is predicted to be slower than for the positive patterning task. This predicted pattern of results is consistent with experimental findings (e.g. Bellingham, Gillette-Bellingham, & Kehoe, 1985). A comparison of the predicted changes in α for the two discriminations reveals that it is predicted to decline to a low level for A and B during positive patterning, and increase to a high level for these stimuli during negative patterning. To our knowledge this prediction has not been tested.

Redhead and Pearce (1993) found that adding a third stimulus to each of the trials of a negative patterning discrimination, to create an AC+ BC+ ABC- discrimination, resulted in

slower acquisition than with an A+ B+ AB- discrimination. The results of a simulation for an AC+ BC+ ABC- discrimination are shown in the bottom row of Figure 7. A comparison of the results shown in Panel A of the bottom two rows indicates that the theory predicts that the acquisition of A+ B+ AB- discrimination is hindered by the presence of a common cue on every trial. A further prediction from the simulation of the AC+ BC+ ABC- discrimination is that C by itself will elicit an abnormally strong response of 2.25λ by the end of training. In support of this prediction, Williams, Mehta, and Dumont (2004; see also Williams, Mehta, Poworczyk, Oriehel, George, & Pearce, 2002) have demonstrated that an AC+ BC+ ABC- discrimination resulted in C acquiring greater than asymptotic associative strength.

The theory thus appears to cope reasonably well with findings from a variety of configural discriminations (see also George and Pearce, 1999). We should, however, draw attention to one finding that poses a challenge to the theory. The simulation of the AC+ BC+ ABC- discrimination also revealed that despite C having considerable associative strength, the value of α for this stimulus is predicted ultimately to be low. In contrast to the prediction, Dopson, Esber, and Pearce (2011) discovered that this training resulted in the salience of C being particularly high. It would be premature to advocate a radical revision of our proposals on the basis of a single result, but this finding indicates a need to investigate further the circumstances that influence changes in attention during the course of configural discriminations. If a comprehensive empirical picture can be drawn of how attention is influenced by such discriminations, we suspect it will provide a particularly demanding test bed for the evaluation of all theories of associative learning, not just the one proposed here.

Concluding comments

In closing this presentation of a configural theory of attention and associative learning, we shall comment briefly on its application to the renewal effect. In an experiment by Bouton

and Peck (1989), rats received conditioning with Stimulus X in Context A and then extinction with X in Context B. Rats were then returned to Context A, where X elicited substantially stronger responding than for a control group that received all three stages in Context A. The renewal of responding in the final stage can be readily explained by the present theory, if it is assumed that during conditioning and extinction the context in which X is presented contributes to the configural representation that enters into an association with the reinforcer. During Stage 1 animals might learn that the configuration AX signals food and during Stage 2 they might learn that BX signals the absence of food. On being returned to Context A for Stage 3, the occurrence of X would then activate the AX configural unit and lead to a recovery of responding. Computer simulations have confirmed this analysis, but they have also revealed that a generalization decrement is expected when X is presented for the first time in Context B. Although there was no hint of such a decrement in the study by Bouton and Peck (1989), it has been observed in other demonstrations of the renewal effect (e.g. Thomas, Larsen and Ayres, 2003). In view of these conflicting findings it may well be worth pursuing further our explanation for the renewal effect.

There is now a strong body of evidence showing that two attentional processes, which we have represented with the parameter α and σ , influence the acquisition of conditioned responding in a variety of contexts. By incorporating these processes into a configural theory of learning it is possible to explain a wide range of experimental findings. As we have seen, however, the theory has its limitations and there remains the challenge of exploring how they might be overcome.

References

- Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. B. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 3, pp. 121-268). *New York: Wiley.*
- Bellingham, W. P., Gillette-Bellingham, K., & Kehoe, E. J. (1985). Summation and configuration in patterning schedules with the rabbit. *Animal Learning and Behavior*, *13*, 152-164.
- Bouton, M. E., & Peck, C. A. (1989). Context effects on conditioning, extinction, and reinstatement in an appetitive conditioning preparation. *Animal Learning and Behavior*, *17*, 188-198.
- Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural Brain Research*, *110*, 67-72.
- Dopson, J. C., Esber, G. R., & Pearce, J. M. (2010). Differences in the associability of relevant and irrelevant stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*, 258-267.
- Dopson, J. C., Esber, G. R., & Pearce, J. M. (2011). Changes in attention to an irrelevant cue that accompanies a negative patterning discrimination. *Learning and Behavior*, *39*.
- Dopson, J. C., Williams, N. A., Esber, G. R., & Pearce, J. M. (2010). Enhanced attention to stimuli that consistently signal the absence of reinforcement. *Learning and Behavior*, *38*, 337-347.
- George, D. N., & Pearce, J. M. (1999). Acquired distinctiveness is controlled by stimulus relevance not correlation with reward. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*, 363-373.

- Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS-US pairings. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 31-42.
- Haselgrove, M., Esber, G. R., Pearce, J. M., & Jones, P. M. (2010). Two kinds of attention in Pavlovian conditioning: Evidence for a hybrid model of learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 456-470.
- Kaye, H., & Pearce. (1984). The strength of the orienting response during Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 90-109.
- LePelley, 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*, 57B, 193-243.
- Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced preexposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*, 52, 415-419.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276-298.
- Mackintosh, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning and Behavior*, 4, 186-192.
- Mackintosh, N. J., & Little, L. (1969). Intradimensional and extradimensional shift learning by pigeons. *Psychonomic Science*, 14, 5-6.
- Pavlov, I. P. (1927). *Conditioned reflexes (G.V.Anrep, Trans.)*. London: Oxford University Press.
- Pearce, J. M. (1987). A model of stimulus generalization for Pavlovian conditioning. *Psychological Review*, 94, 61-73.

- Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, 101, 587-607.
- Pearce, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning and Behavior*, 30, 73-95.
- Pearce, J. M., Esber, G. R., George, D. N., & Haselgrove, M. (2008). The nature of discrimination learning in pigeons. *Learning and Behavior*, 36, 188-199.
- Pearce, J. M., George, D. N., & Redhead, E. S. (1998). The role of attention in the solution of conditional discriminations. In N. A. Schmajuk & P. C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals* (pp. 249-275). Washington, DC: APA Press.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not 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*. Oxford: Oxford University Press.
- Pearce, J. M., & Redhead, E. S. (1993). The influence of an irrelevant stimulus on two discriminations. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 180-190.
- Reberg, D., & LeClerc, R. (1977). A feature positive effect in conditioned suppression. *Animal Learning and Behavior*, 5, 143-147.
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, 72, 77-94.
- Rescorla, R. A. (2000). Associative change in exciters and inhibitors differ when they are conditioned in compound. *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 428-438.

- Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.
- Swan, J. A., & Pearce, J. M. (1988). The orienting response as an index of stimulus associability in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 292-301.
- Swartzentruber, D., & Bouton, M. E. (1986). Contextual control of negative transfer produced by prior CS-US pairings. *Learning and Motivation*, *17*, 366-385.
- Thomas, B. L., Larsen, N., & Ayres, J. J. B. (2003). Role of context similarity in ABA, ABC, and AAB renewal paradigms: Implications for theories of renewal and for treating human phobias. *Learning and Motivation*, *34*, 410-436.
- Williams, D. A., Mehta, R., & Dumont, J. L. (2004). Conditions favoring superconditioning of irrelevant conditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 148-159.
- Williams, D. A., Mehta, R., Poworonyk, T. M., Oriehel, J. S., George, D. N., & Pearce, J. M. (2002). Acquisition of superexcitatory properties by an irrelevant background stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*, 284-297.
- Wilson, P. N., Boumphrey, P., & Pearce, J. M. (1992). Restoration of the orienting response to a light by a change in its predictive accuracy. *Quarterly Journal of Experimental Psychology*, *44B*, 17-36.

Table 1. Design of an IDS/EDS experiment with A, B, C and D belonging to one dimension, and W, X, Y and Z belonging to another dimension.

Group	Stage 1	Stage2
IDS	AW+ BW- AX+ BX-	CY+ DY- CZ+ DZ-
EDS	AW+ BW- AX+ BX-	CY+ CZ- DY+ DZ-

Footnote to Title Page

A MATLAB simulator of the connectionist network described here can be downloaded from <http://psy.hull.ac.uk/Resources/Immediacy/georgepearce.zip>. This simulator is compatible with MATLAB R2009b and later.

This work was supported by a Royal Society University Research Fellowship to DNG and a grant from the UK Biotechnology and Biological Sciences Research Council to JMP.

David N George, Department of Psychology, University of Hull, Hull, HU6 7RX, United Kingdom; and School of Psychology, University of New South Wales, Sydney, Australia. E-mail: D.George@Hull.ac.uk

John M Pearce, School of Psychology, Cardiff University, Cardiff, CF10 3AT, United Kingdom. E-mail: pearcejm@cardiff.ac.uk

Figure Captions

Figure 1. Connectionist network for a configural theory of associative learning showing the connections involved in the solution of an AX+ BX- discrimination. The feedback links permit changes in the conditionability of configural units (dotted lines), and changes in the relative effectiveness of input units (dashed lines).

Figure 2. The predicted changes in, *Panel A*, the strength of response to a pattern of stimulation; *Panel B*, the associative strength of the configural unit activated by the pattern; *Panel C*, the effective salience, α , of the cues creating the pattern; and *Panel D*, the associability, σ , of the configural unit activated by the pattern. In Row 1 (top row), the pattern was a single stimulus, A, paired with a weak US. In Row 2 the pattern was the same stimulus, or a novel stimulus, B, paired with a large US. In Row 3 the pattern was a light paired with a tone-food sequence on either a continuous (Group C) or partial (Group P) reinforcement schedule. In Row 4 the pattern was the light from Row 3 paired directly with food.

Figure 3. *Panels A* through *D* have the same significance as for Figure 2. In Row 1 (top row) the pattern was a compound, AB, of two equally salient stimuli paired with a US. In Row 2 the pattern was a compound, AB, of two stimuli paired with a US with A being more salient than B. In Row 3 the pattern was a compound, AB, of two equally salient stimuli paired with a US after A had been paired with the same US.

Figure 4. The predicted changes in, *Panel A*, the strength of responding to the two trials of a feature-negative, A+ AB-, and a feature-positive, C- CD+ discrimination; *Panel B*, the

associative strengths of the configural units activated by the four different types of trial of the discriminations considered in Panel A; *Panel C*, the effective salience, α , of the four stimuli from the discriminations considered in Panel A; and *Panel D*, the associability, σ , of the configural units considered in Panel B.

Figure 5. The predicted changes in, *Panel A*, the strength of responding to the two trials of an AX+ BX- discrimination; *Panel B*, the associative strengths of the configural units activated by the two different types of trial of the discriminations considered in Panel A; *Panel C*, the effective salience, α , of the three stimuli from the discriminations considered in Panel A; and *Panel D*, the associability, σ , of the configural units considered in Panel B.

Figure 6. Panels A through D of the top row have the same significance as for Figure 2, but the patterns were from an IDS/EDS experiment. In Panel C relevant refers to unique features of A and B, and irrelevant to common features of A and B and all features of X and Y. The panels in Row 2 depict changes in the strength of responding to new patterns during subsequent discrimination training with an IDS group and an EDS group.

Figure 7. Panels A through D have the same significance as for Figure 2. In Row 1(top row) the patterns were used for A- B- AB+ discrimination. In Row 2 the patterns were from an A+ B+ AB- discrimination. In Row 3 the patterns were from an AC+ BC+ ABC- discrimination.

Figure 1

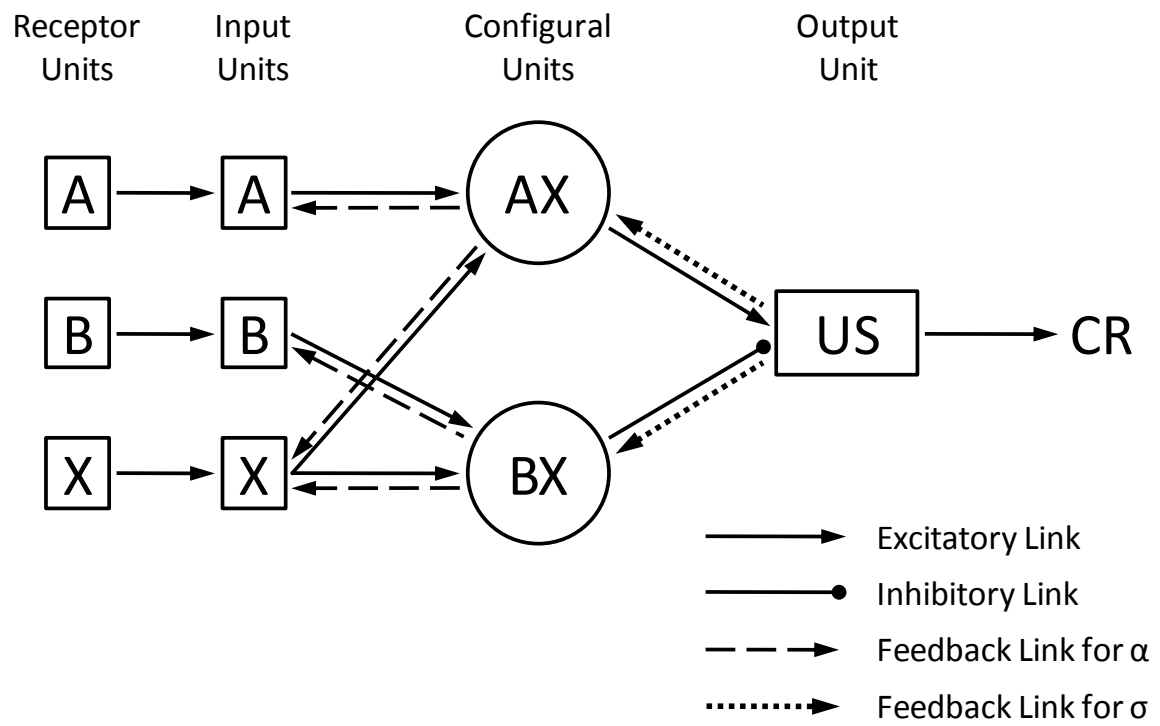


Figure 2

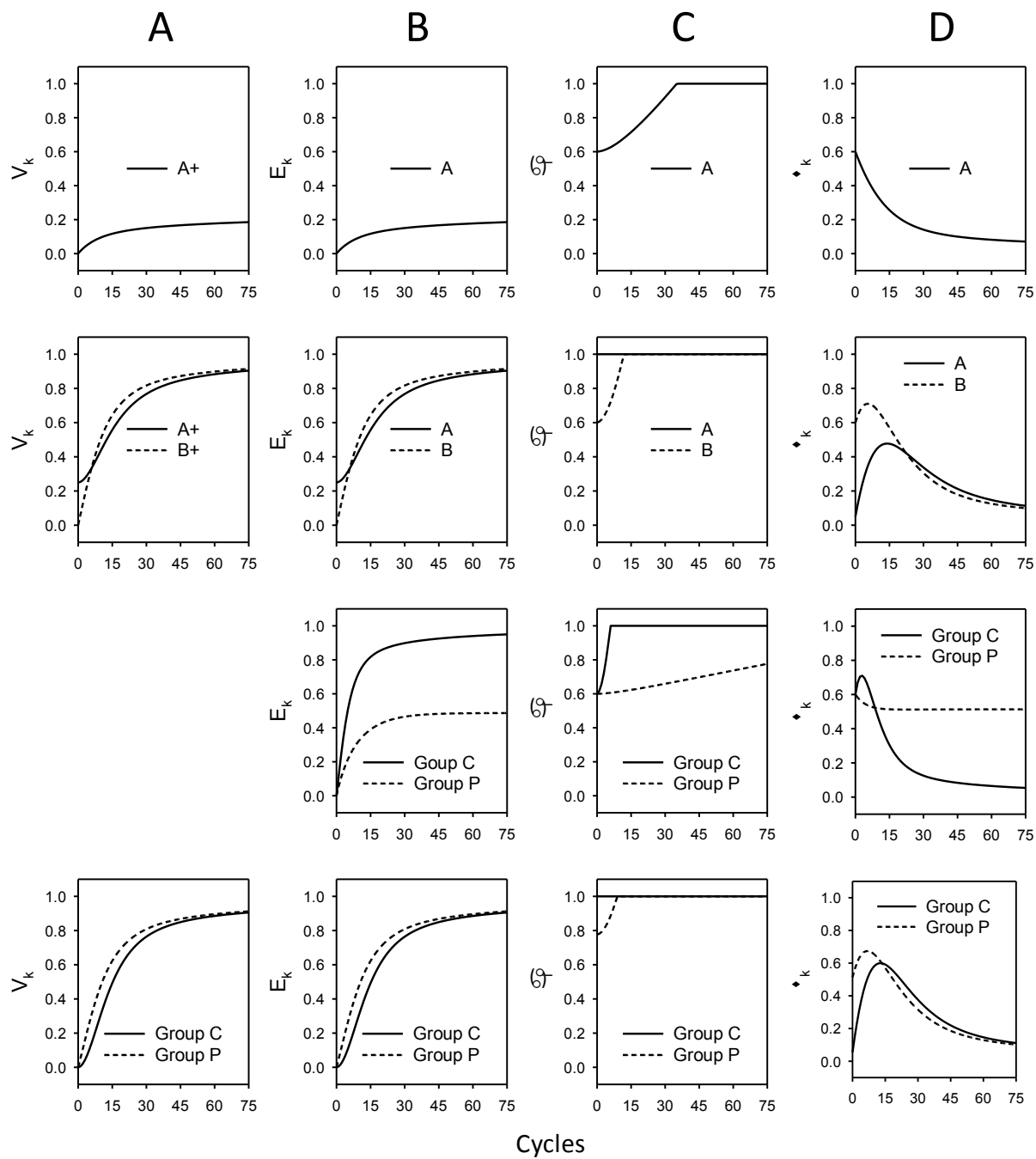


Figure 3

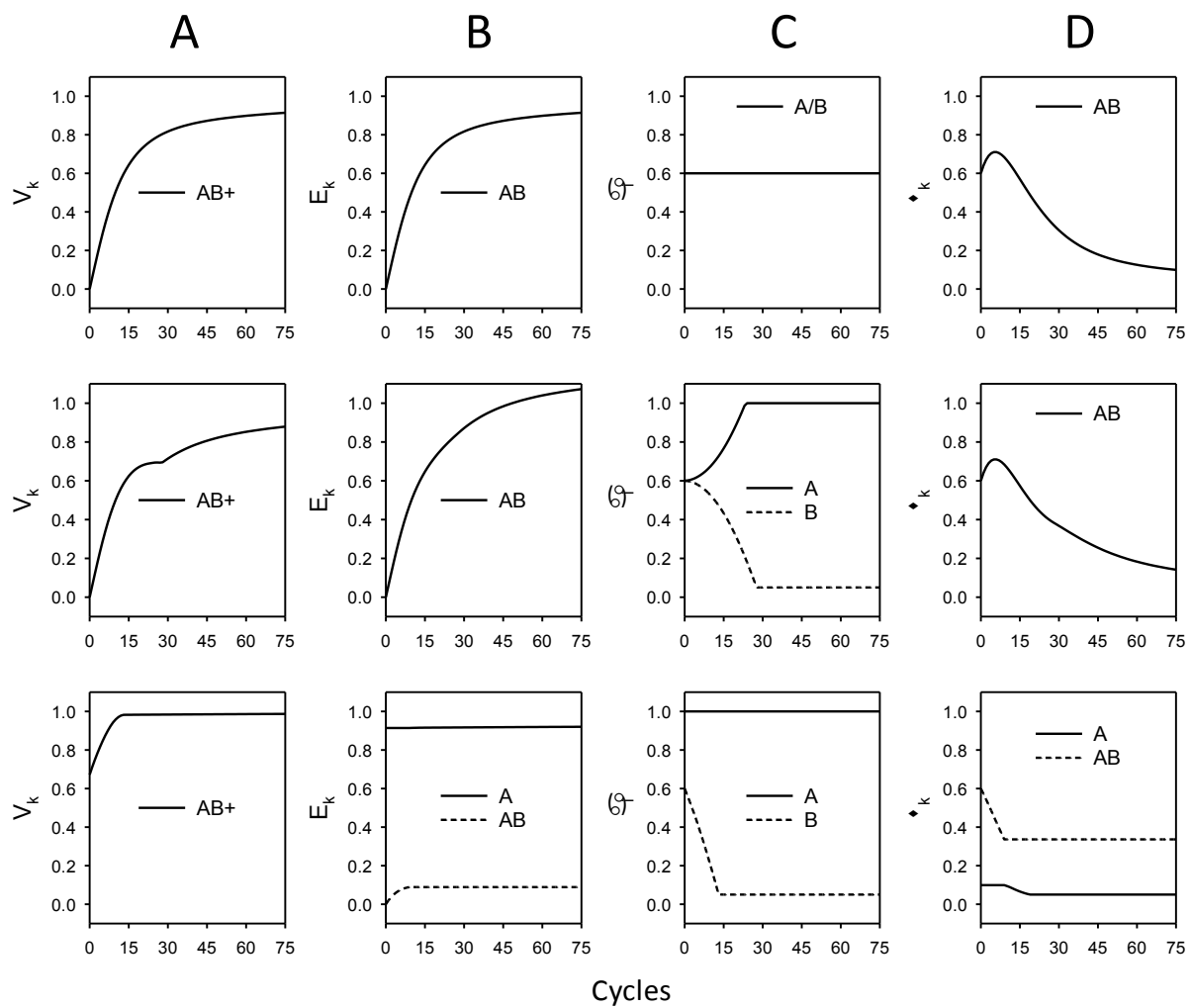


Figure 4

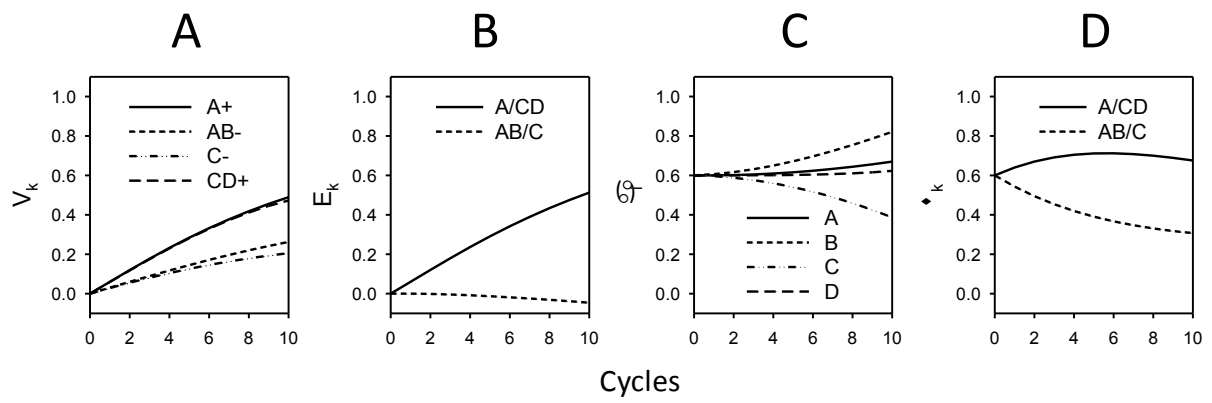


Figure 5

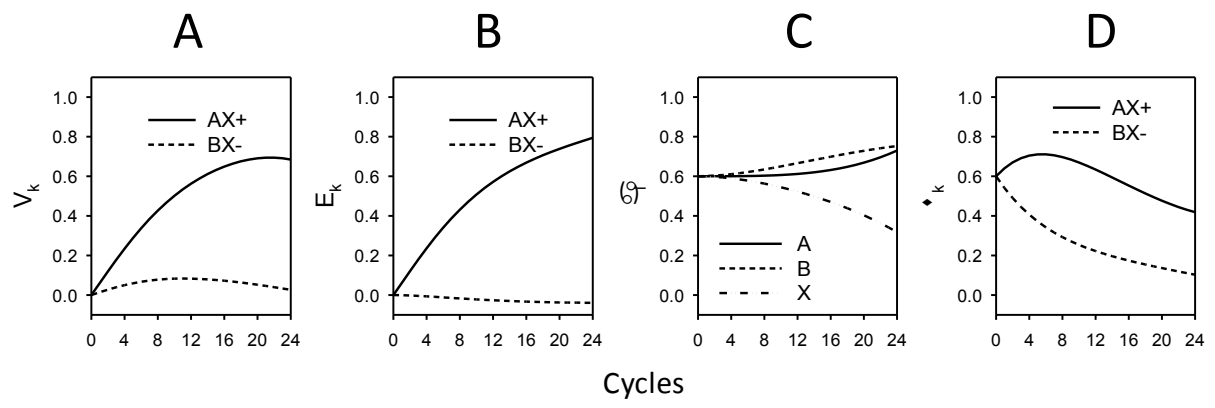


Figure 6

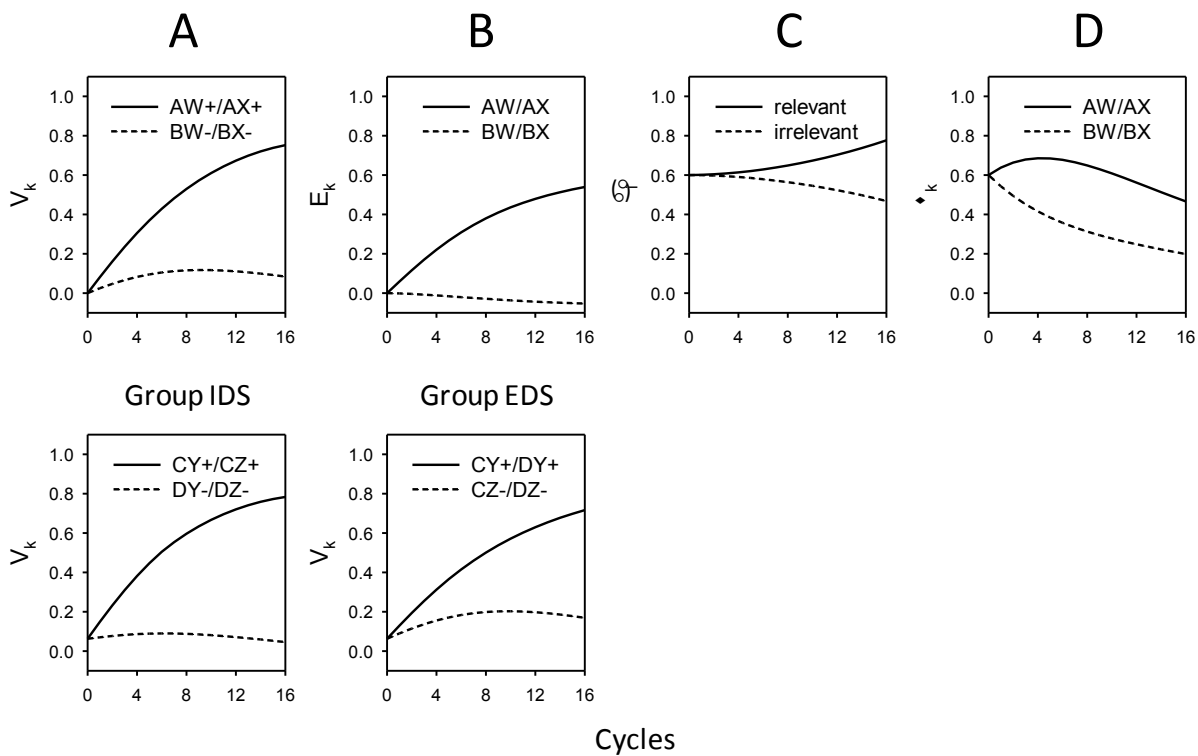


Figure 7

