The nature of discrimination learning in pigeons

JOHN M. PEARCE, GUILLEM R. ESBER, AND DAVID N. GEORGE
Cardiff University, Cardiff, Wales

AND

MARK HASELGROVE
University of Nottingham, Nottingham, England

The results from five experiments are considered in relation to two of Spence's (1937, 1938) proposals concerning discrimination learning. In Experiments 1 and 2, we investigated whether his ideas about the interaction between excitatory and inhibitory generalization gradients can be used to understand how animals solve a complex patterning discrimination. The results supported a development of his proposals as put forward by Pearce (1994), provided a modification was made to Pearce's rule for determining the shape of the generalization gradient. In Experiments 3, 4, and 5, we examined whether animals would pay more attention to stimuli that are relevant, rather than irrelevant, to the solution of a discrimination. The results supported this proposal for stimuli comprising visual patterns, but not for those comprising plain colors. The results also indicated that change of attention was a consequence of preliminary receptor-exposure acts, as envisaged by Spence, and not of more central changes in attention.

Although it has been more than 70 years since Spence (1936) published his classic article, "The Nature of Discrimination Learning in Animals," the ideas in it and in his subsequent articles (Spence, 1937, 1938, 1940, 1942, 1952) remain influential. For example, he rejected the noncontinuity view that during a discrimination, animals try out a succession of hypotheses about which stimulus will signal a reward until they discover the correct one and the problem is solved (e.g., Krechevsky, 1932; Lashley, 1929). Instead, he advocated a continuity position, whereby learning at the outset of a discrimination was assumed to progress incrementally with all the stimuli that were perceived (Spence, 1940). The importance of this proposal is apparent, because it has been incorporated into the majority of currently influential theories of associative learning (e.g., Rescorla & Wagner, 1972). Another enduring influence of the theory has been the proposal that after subjects have received a discrimination between two stimuli from the same dimension, responding to those stimuli, as well as to others from the same dimension, is determined by the interaction between excitatory and inhibitory generalization gradients (see Figure 1). The importance of this proposal is apparent: The majority of current theories assume that the solution of many discriminations involves the acquisition of excitatory and inhibitory associative strength. Furthermore, they assume that the algebraic interaction between these opposing strengths will determine the magnitude of a conditioned response (e.g., Pearce, 1994; Rescorla & Wagner, 1972).

Other aspects of Spence's theory have not withstood the test of time. The theory has been described as nonselective

because the increment in associative strength to a stimulus on a trial is assumed to be unaffected by the associative properties of the stimuli that accompany it. This assumption leads to a number of incorrect predictions about a variety of experimental findings. One such finding is the *relative-validity effect*, which was first demonstrated by Wagner, Logan, Haberlandt, and Price (1968). An experimental group received a true discrimination, AX+/BX−, in which a compound of two stimuli signaled reward (AX+), and a compound of two stimuli signaled the absence of reward (BX−). A control group received a pseudodiscrimination, AX±/BX±, in which both compounds signaled food on half the trials. Stimulus X was thus paired with food according to the same intermittent reinforcement schedule for both groups. It follows from the nonselective principle that these different treatments will result in similar rates of responding during X, if it should subsequently be presented by itself. In contrast to this prediction, Wagner et al. discovered that the rate of responding during a test with X was stronger after training with the pseudo- than with the true discrimination.

A variety of explanations have been developed for the relative-validity effect, and they reveal the different directions that theorizing about the nature of discrimination learning in animals has taken since Spence put forward his ideas. Some theories have replaced Spence's nonselective principle with the idea that the growth of associative strength by a stimulus is affected by the associative properties of the stimuli that accompany it. The most influential theory to make this assumption was proposed by

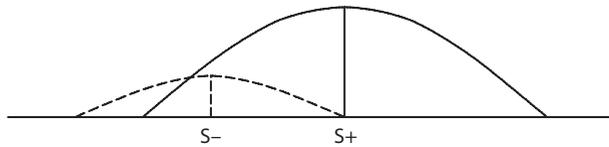**J. M. Pearce, pearcejm@cardiff.ac.uk**

Figure 1. The gradients of excitation (solid curve) and of inhibition (dashed curve) that are predicted by Spence's (1937) theory to develop during a discrimination in which reward is presented in the presence of one stimulus (S+), but not of another (S−). The strength of responding is determined by subtracting the inhibitory curve from the excitatory curve.

Rescorla and Wagner (1972). According to this theory, the increment in associative strength on any trial to an individual stimulus is determined by the discrepancy between the sum of the associative strengths of the stimuli that are present and an asymptotic value determined by the magnitude of the reinforcer. Wagner and Rescorla (1972) showed how the relative-validity effect can be explained by this type of theory. Another class of explanation has retained Spence's idea that changes in associative strength to a stimulus on a given trial are unaffected by the associative properties of the stimuli that accompany it, but assumed that the attention paid to stimuli varies during the course of a discrimination (e.g., Sutherland & Mackintosh, 1971; Mackintosh, 1975). Thus, for a true discrimination, X is less relevant than A or B to its solution, and Mackintosh's (1975) theory predicts that attention to X will decline. As a result, even though the associative strength of X can change on every trial, the changes will be slight because of the low attention it receives. In contrast, for the pseudo-discrimination, X is more relevant than A and B as a signal for food. Attention to X is predicted to be relatively high, which will then allow it to gain more associative strength than in the true discrimination.

These two explanations can be regarded as being elemental because they share the assumption with Spence (1936) that at the outset of a discrimination, each stimulus has the opportunity to enter into an association with the trial outcome. Yet another explanation for the relative-validity effect can be based on the assumption that on any trial, only a single association will develop, which will be between the entire pattern of stimulation on that trial and the trial outcome (Pearce, 1987, 1994). A convenient way to understand this explanation for the relative-validity effect is to return to the conceptualization shown in Figure 1 of how animals solve discriminations. Instead of the abscissa representing different values from the same physical

dimension, it could represent values from a more abstract dimension, which represents the similarity between different patterns of stimulation. Figure 2 shows this modification for the true discrimination used by Wagner et al. (1968). Compounds AX and BX have been located on the similarity dimension, with an excitatory gradient centered on AX and an inhibitory gradient centered on BX. Also located on the abscissa is the point occupied by X when it is presented alone. Even though the figure shows that more excitation will generalize to X from AX than will inhibition from BX, the difference between the two gradients will ensure that the response to X will be relatively weak. On the other hand, a pseudodiscrimination will result in both AX and BX acquiring excitatory strength, with the result that the net level of excitation that generalizes to X will be considerably greater than the net level of excitation after the true discrimination.

It is not possible in a single article to provide a comprehensive evaluation of these different successors to Spence's theory. Rather, in the hope of providing an indication of how Spence's ideas are still relevant to the analysis of discrimination learning, the present article will concentrate on two questions raised by the foregoing discussion. First, what is the role of configural information in animal discrimination learning and, second, what is the role of attentional processes?

A Configural Explanation for Stimulus Generalization

If a configural explanation for how animals solve discriminations is to have any merit, it must formally specify how the similarity between two patterns of stimulation is determined. Once this measure has been derived, it will then be possible to make precise predictions about the extent of generalization between the patterns. On the basis of principles embodied in his connectionist network for discrimination learning, Pearce (1994) initially proposed that the similarity between patterns A and B—$_AS_B$—is determined by Equation 1, where n_A and n_B represent the number of elements in A and B (respectively), and n_C represents the number of elements that are common to both patterns:

$$_AS_B = \frac{n_C}{\sqrt{n_A \times n_B}}. \tag{1}$$

Once this expression is incorporated into the equations described below, it can be used to predict the course of acquisition of a discrimination between the two patterns. Although these equations function well for simple discriminations, they incorrectly predict that animals will be unable to solve a negative patterning discrimination, A+/B+/AB− (e.g., Woodbury, 1943). In fact, because of excessive excitatory generalization from the components to the compound, the network predicts that responding to AB will be consistently stronger than to A or B alone. One solution to this problem is to reduce the generalization between patterns by reducing the degree of similarity between them. Because the value of $_AS_B$ varies between 0
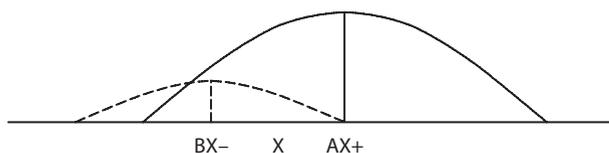

Figure 2. One way in which Spence's (1937) theory could be developed in order to provide a configural explanation for the solution of a true discrimination, AX+/BX−, and to predict the strength of responding that will occur on a test trial with X alone.

and 1, a simple method for reducing the similarity between two patterns is to raise the expression in the right-hand side of Equation 1 to the power of 2, as shown in Equation 2. Once this change is made, Equation 2, together with Equations 3 and 4, is able to predict that a negative patterning discrimination will be solved (see Pearce, 1994):

$$_AS_B = \left(\frac{n_C}{\sqrt{n_A \times n_B}} \right)^2. \qquad (2)$$

A benefit of this modification, which was advocated by Pearce (1994), is that the new measure of similarity is equivalent to measures proposed by Pearce (1987) and Atkinson and Estes (1963). Equation 2 also captures the principle that the similarity between two patterns is related directly to the number of elements they have in common. Equation 3 was proposed by Pearce (1994) in order to describe how the associative strength of pattern B, V_B, will generalize to pattern A, E_A, where the value of $_AS_B$ is given by Equation 2.

$$E_A = {}_AS_B * V_B. \qquad (3)$$

Finally, Pearce (1994) proposed that the increment in associative strength to a pattern of stimulation on any trial, ΔV_A, is given by Equation 4, where β is a learning rate parameter ($0 < \beta < 1$), λ is the asymptote of conditioning,
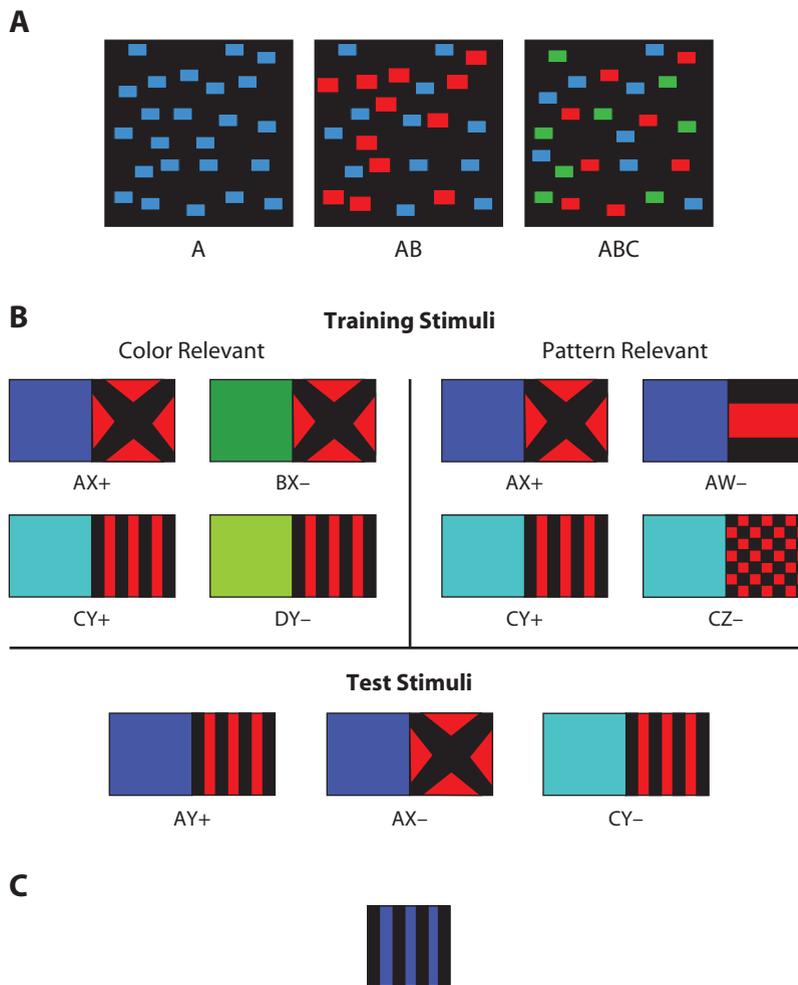

Figure 3. (A) The stimuli that were used for the trials with A, AB, and ABC in Experiment 1 (there were 120 rectangles in each pattern). (B) The training stimuli above the horizontal line were used for Experiment 3. For the color-relevant group, colors A, B, C, and D were relevant to the trial outcome, and patterns W, X, Y, and Z were irrelevant; for the pattern-relevant group, patterns W, X, Y, and Z were relevant to the trial outcome, and colors A, B, C, and D were irrelevant. The test stimuli below the horizontal line were given to the two groups for the test trials of Experiment 3. Assuming that more attention is paid to stimuli that were relevant rather than irrelevant during the training stage, the discrimination between AY and CY will be easy for the color-relevant group and hard for the pattern-relevant group; the discrimination between AY and AX will be easy for the pattern-relevant group and hard for the pattern-relevant group. (C) Example of an integrated stimulus for Experiment 4.

and the expression within the inner pair of parentheses represents the associative strength of A plus the associative strength that generalizes to A as a result of conditioning with other stimuli. In keeping with other selective theories (e.g., Rescorla & Wagner, 1972), it follows from Equation 4 that any increase in associative strength to one stimulus will be influenced by the associative properties of the stimuli that accompany it:

$$\Delta V_A = \beta \times [\lambda - (V_A + E_A)]. \qquad (4)$$

An advantage of these proposals is that they provide a ready account for the influence of similarity on the ease with which a discrimination is solved (Pearce, 1994), and they are also able to explain a good number of findings from studies of Pavlovian conditioning (see Pearce, 2002, for a review). There is, however, one prediction from configural theory that has never been tested and appeared to us unlikely to be confirmed. In order to further evaluate a configural analysis of discrimination learning, therefore, we tested this prediction in two experiments.

The novel prediction concerns a complex patterning discrimination in which three stimuli are presented in all possible combinations, and is closely related to a study by Redhead and Pearce (1995). In that experiment, pigeons received autoshaping with numerous dots of three different colors, A, B, and C, presented on a television screen behind a response key (see panel A of Figure 3). Trials with dots of a single color, A+/B+/C+, or of two colors, AB+/AC+/BC+, were always followed by food, whereas trials with dots of three colors on the screen, ABC−, were never followed by food. The subjects solved the discrimination, and as they did so, responding was more vigorous on the trials with individual colors than on those with pairs of colors. According to configural theory, this outcome makes sense because the nonreinforced trials with ABC will result in its acquiring inhibition. As ABC acquires inhibition, it disrupts responding to the reinforced stimuli through the effects of generalization. The disruptive effects of this generalization will, however, be more marked on the trials with the pairs of stimuli because of their close similarity with ABC than on trials with individual stimuli because of their more distant similarity with ABC. When these principles are applied to the discrimination to be considered next, they lead to an unexpected prediction.

Experiment 1

Pigeons were trained, using autoshaping, with the stimuli just described for an A+B+C+/AB−AC−BC−/ ABC+ patterning discrimination in which A, B, C, and ABC signaled food, and in which AB, AC, and BC were followed by nothing. The interaction between excitatory and inhibitory generalization gradients with this arrangement is complex, and in order to determine the prediction made by configural theory concerning the proposed design, a computer simulation based on Equations 2, 3, and 4 was conducted. In keeping with the experimental design, each cycle of the simulation contained a single trial, with each of the seven possible patterns presented in a random sequence. The value of β was .2, and the value of λ was 1.0.
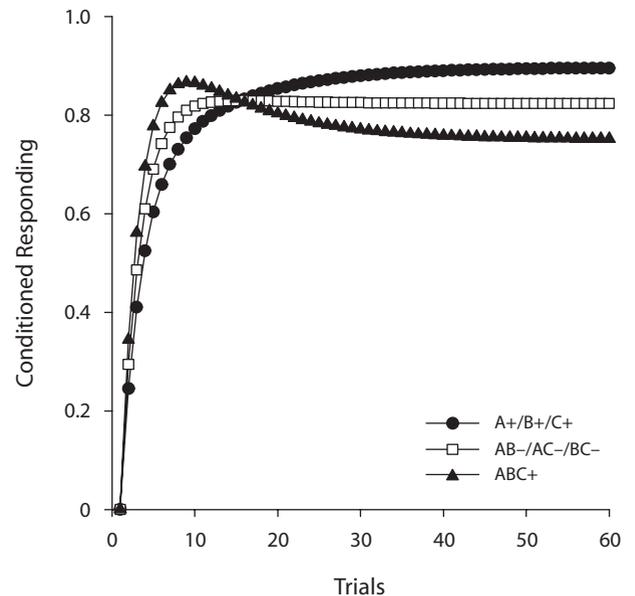

Figure 4. A computer simulation based on Equations 2, 3, and 4 to determine the predictions made by Pearce's (1994) theory concerning the outcome of an A+B+C+/AB−AC−BC−/ABC+ discrimination.

The results from the simulation are presented in Figure 4, which shows that responding is predicted to be fastest to A, B, and C, slowest to ABC, and at an intermediate rate to AB, AC, and BC. In other words, the theory predicts that animals will fail to solve the discrimination, and they will fail to solve it in a particular way. Our expectation was that the pigeons would solve the discrimination, and if this proved to be correct, the theory would need modification.

A single group of pigeons received autoshaping, in which a training pattern was presented for 10 sec on a television screen and, where appropriate, was followed by the presentation of food for 4 sec. The screen was located behind a clear Perspex panel that was attached to the test chamber by a hinge at the top. The training pattern for every trial consisted of 120 small colored rectangles that were randomly located within a matrix that could accommodate 20 rectangles horizontally and 32 rectangles vertically. The colors of A, B, and C were red, green, and blue, respectively. When more than one color was presented on the screen, there were an equal number of dots of each color. Following preliminary autoshaping, with a white circle signaling food, the discrimination was introduced with each of the seven patterns shown eight times in a session. The results from the experiment are shown in Figure 5, where it can be seen that the A+B+C+/AB−AC−BC−/ABC+ discrimination was solved. Details of the training that we have omitted can be found in Redhead and Pearce (1995).

The problem with configural theory, as far as the present results are concerned, is that it predicts that there will be too much generalization among the patterns. The discrimination will result in patterns A, B, C, and ABC entering into excitatory associations, which will all generalize to AB, AC, and BC and result in each pair of stimuli arousing a high level of generalized excitation. In order
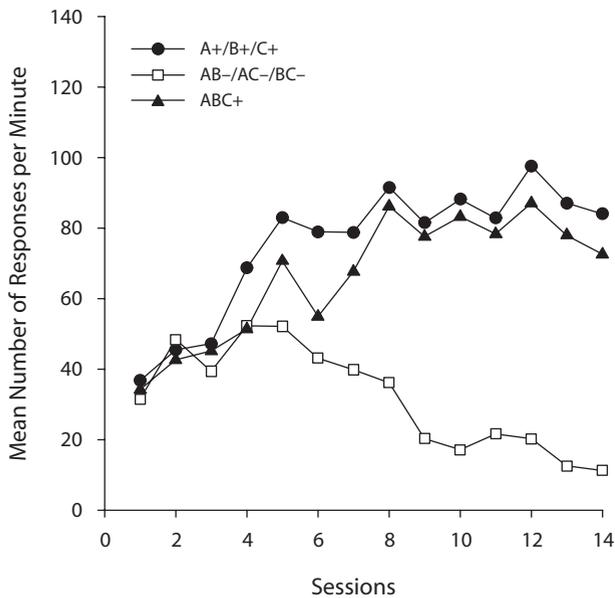
Figure 5. The mean number of responses per minute by a group of pigeons to the three different types of trials of an A+B+C+/ AB−AC−BC−/ABC+ discrimination during the 14 sessions of Experiment 1.

to counter this excitation, AB, AC, and BC will enter into inhibitory associations which, once developed, will generalize to ABC and to A, B, and C, and will weaken responding in their presence. To compensate for this effect, it will be necessary for the excitatory associations to gain even more strength, and through generalization, they will continue to support a high rate of responding to AB, AC, and
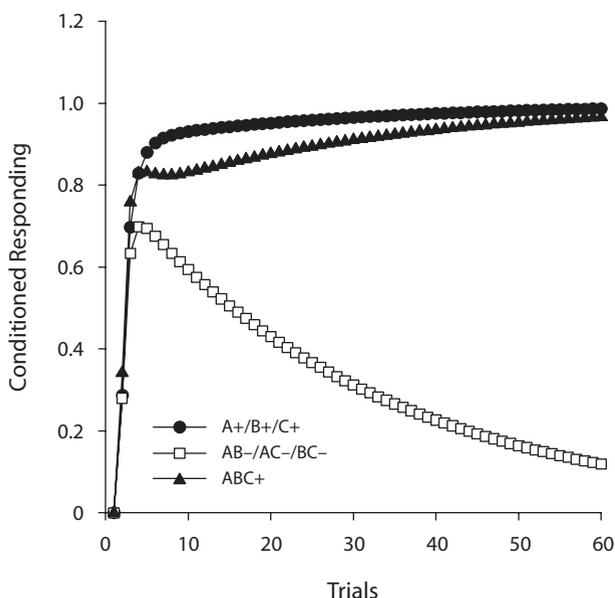

Figure 6. A computer simulation based on Equations 2, 3, and 4 (with the exponent in Equation 2 raised to 3) to determine the predictions made by Pearce's (1994) theory concerning the outcome of an A+B+C+/AB−AC−BC−/ABC+ discrimination.

BC. Thus, despite extended nonreinforcement, configural theory predicts—incorrectly, it turns out—that responding to the pairs of stimuli will be sustained at a high and faster rate than to ABC.

If a way could be found to reduce the degree of predicted generalization among the different types of compounds, it is conceivable that configural theory would be able to explain the results from Experiment 1. A simple method for achieving this goal is to follow the strategy previously adopted and again raise the value of the exponent for the measure of similarity, this time from 2 to 3 (e.g., Kinder & Lachnit, 2003). In addition to sharpening generalization gradients around individual patterns, as far as we can tell, this modification does not reverse any of the successful predictions of the theory. The modification does, however, alter the prediction concerning the experiment that has just been described. The results from a further computer simulation for Experiment 1, but with the exponent in Equation 1 raised to 3, can be seen in Figure 6. The fit between the predicted pattern of results and the results shown in Figure 5 is, to our eyes, reassuring.

Experiment 2

An implication of the foregoing analysis is that the outcome of an A+B+C+/AB−AC−BC−/ABC+ discrimination is determined by the amount of generalization among the various patterns. If the degree of generalization is moderate, configural theory predicts that animals will be able to solve the discrimination. However, as revealed by the first computer simulation, the theory predicts that if the level of generalization among the patterns is considerable, animals will fail to solve the discrimination, and they will fail to solve it in a particular way. Experiment 2 was conducted in order to test this prediction. A single group of pigeons received an A+B+C+/AB−AC−BC−/ABC+ discrimination, but the stimuli were different from those used for the previous experiment. The three stimuli, A, B, and C, were single white circles with a diameter of 1 cm and were presented on the television screen at the three corners of a notional equilateral triangle (with sides of 2.4 cm). Sketches of three of the experimental stimuli can be seen in Figure 7.

For trials with compounds composed of pairs of stimuli, two circles were presented on the television screens, and for trials with ABC, three circles were shown. We had discovered from a preliminary study that after conditioning with ABC by itself, there was a substantial degree of generalization to the pairs of circles, and also to individual circles. The same study also revealed that after conditioning with ABC, using the colored rectangles of the previous experiment, the degree of generalization to the pairs of colors and to the individual colors was significantly less than when the white circles served as A, B, and C. Thus, it is reasonable to assume that the degree of generalization among the training patterns for Experiment 2 was greater than that for Experiment 1, and it is reasonable to hope, therefore, that the outcome of the A+B+C+/ AB−AC−BC−/ABC+ discrimination will be more in keeping with the predicted pattern of results shown in Figure 4 than that shown in Figure 6. Except for the use
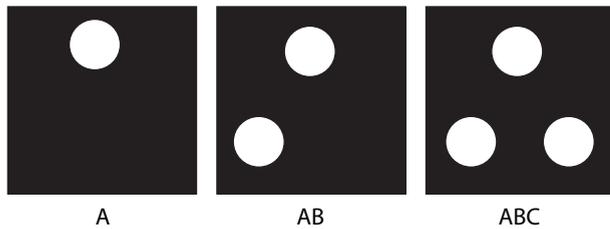
Figure 7. The stimuli that were used for trials with A, AB, and ABC in Experiment 2. They were composed of white circles against a black background.

of different stimuli, the procedure was the same as for Experiment 1. The results from the experiment are shown in Figure 8. The pattern of responding during the three types of trials was remarkably similar to that shown in the computer simulation in Figure 3. In other words, taking steps to increase generalization among the patterns prevented the pigeons from solving the discrimination. They did not, however, treat the patterns indiscriminately; they responded to them in an orderly manner that was predicted on the basis of considerations about the interaction between excitatory and inhibitory generalization gradients.

Discussion

We have already noted that Spence (1936) offered an elemental analysis of discrimination learning, but this does not mean that he completely ignored the role of configural information. Spence (1952) was aware that animals can solve patterning discriminations in which the outcome is determined by the significance of combinations of stimuli, rather than by individual stimuli, and he appreciated that these could not be solved by associations
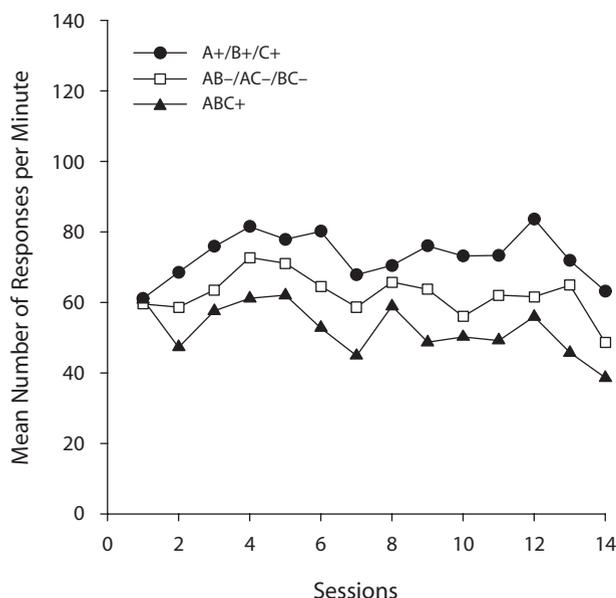

Figure 8. The mean number of responses per minute by a group of pigeons to the three different types of trials of an A+B+C+/ AB−AC−BC−/ABC+ discrimination during the 14 sessions of Experiment 2.

based on individual experimental cues. To account for the successful solution of this type of discrimination, he made use of Hull's (1943) afferent interaction hypothesis. For instance, to explain the solution of a simple negative patterning discrimination (A+B+/AB−), the stimuli would be assumed to be perceived differently when they were presented separately relative to when they were presented together. As a consequence, this discrimination would effectively involve four stimuli—a and b, which would be present during the A+ and B+ trials, respectively, and a' and b', which would both be present during the AB+ trials. The discrimination would be solved through a and b's gaining excitatory strength and a' and b''s gaining inhibitory strength. A more sophisticated development of these ideas has been advanced by Wagner (2003, 2008), who suggested that any stimulus is composed of a number of elements, each of which may enter into excitatory and inhibitory associations. When two stimuli are presented together, some of the elements they each excite when presented alone are held to be replaced by elements that are excited only when the stimuli are presented together. Thus, the representation of a stimulus when it is presented alone is rather different from that when it is accompanied by another stimulus; to this extent, Wagner's (2003, 2008) theory can be regarded as providing a mechanism for accounting for Hull's afferent interaction hypothesis.

In order to determine whether these ideas can explain the results from the first two experiments, a series of computer simulations was conducted based on the principles advocated by Wagner (2003). Over the wide range of parameter values that we selected, we consistently found that the theory predicts that the A+B+C+/AB−AC−BC−/ ABC+ discrimination will result in responding during the nonreinforced trials eventually being slower than that during the reinforced trials. Thus, the results from Experiment 2 appear to confound this theory. Furthermore, in all of the simulations, responding was predicted to be consistently more vigorous to the triple element compound, ABC, than to the elements by themselves. Inspection of panel A of Figure 3 and of Figure 5 reveals that we found the opposite outcome in both experiments, but it was significant only in Experiment 2; this appears to pose a further problem for the theory. We would not want to argue that it is impossible to explain the present results with an elemental theory of learning. A subtle modification of the proposals of Wagner (2003, 2008), or to some other elemental theory of learning (e.g., McLaren & Mackintosh, 2000, 2002), may well explain our findings. Indeed, a great benefit of having opposing theories of learning is that the tension between them will lead to their being refined in order to accommodate the findings from experiments that were designed to choose among them.

A particularly stringent test of any theory is for it to correctly predict a novel result that does not accord with common sense. The present experiments were conducted because we considered the prediction shown in Figure 2 to be so implausible that a test of it would reveal a shortcoming of configural theory. The fact that the theory was correct, at least in certain conditions, suggests that there may be some merit in further pursuing this particular approach.

The Role of Attention
in Discrimination Learning

The rationale for the next set of experiments can be appreciated by considering the true discrimination that was referred to at the beginning of this article—AX+/BX−. Cues A and B are relevant to the solution of the discrimination by virtue of signaling, respectively, the occurrence or omission of food, and cue X is irrelevant because it occurs on both types of trials. For over 60 years, it has been suggested that when animals are confronted with a true discrimination, they may pay more attention to the cues that are relevant than those that are irrelevant to its solution (e.g., Lawrence, 1949, 1950).

Tests of this proposal have a checkered history. Early experiments that were said to support this idea (Lawrence, 1949) were subsequently criticized for being open to alternative interpretations (Siegel, 1969; but see Sutherland & Mackintosh, 1971). More recent experiments have compared the effects of training subjects on a true discrimination involving stimuli from two dimensions—with one dimension being relevant and one irrelevant—and then giving them a new discrimination with new stimuli from the same dimensions. In general, the new discrimination is acquired more readily if the dimension that is relevant for the first discrimination is relevant for the second one (intradimensional shift) than if the originally irrelevant dimension becomes relevant (extradimensional shift). Although these results have been taken to be evidence that animals pay more attention to the relevant than to the irrelevant dimension during the first stage (e.g., Mackintosh, 1974; Sutherland & Mackintosh, 1971), Hall (1991) has advanced an alternative explanation in terms of acquired equivalence. The relevant cues in a true discrimination signal different outcomes, whereas the irrelevant cues signal the same outcome. It is thus possible that members of the irrelevant dimension are treated as being equivalent and hence more difficult to differentiate than those belonging to the relevant dimension. Hall (p. 206) admits that this argument is "exceedingly tentative," but it serves to emphasize that it is open to question whether animals pay more attention to relevant than to irrelevant stimuli as a consequence of being exposed to a true discrimination. In an attempt to gain a clearer understanding of the role that attentional processes play in discrimination learning, we have initiated a series of experiments that adopt a new method for investigating the changes in attention that take place during a true discrimination.

Experiment 3

The subjects were again pigeons, the method of training was again autoshaping, and the apparatus was the same as that used for the previous experiments. The experimental stimuli consisted of two adjacent squares with sides of 2 cm. They were located in the middle of the television screen, which was behind the Perspex response key. One square was always filled with one of four plain colors, and the other was filled with one of four patterns. On half the trials, the color was in the left-hand square, and on the remaining trials, it was in the other square. Throughout this discussion, the letters "A" through "D" refer to colors, and the letters "W" through "Z" refer to patterns. The training that was given with some of these stimuli to the two groups of Experiment 3 is summarized in panel B of Figure 3. Panel B shows that the color-relevant group received two true discriminations, with four colors serving as the relevant cues signaling the presence and absence of food, and two patterns serving as the irrelevant cues: AX+/BX− and CY+/DY−. The pattern-relevant group was exposed to similar stimuli, but this time, four patterns were relevant and two colors irrelevant: AX+/AW− and CY+/CZ−. In fact, the training was more complex than that shown in panel B of Figure 3, in order to ensure that each group received the same amount of exposure to the relevant and irrelevant stimuli. Thus, the color-relevant group received additional training with AW+/BW− and CZ+/DZ−, whereas the pattern-relevant group received additional training with BX+/BW− and DY+/DZ−. Each of the eight different trials occurred 10 times in each of 32 sessions of training in Stage 1.

Once the discriminations had been solved, the birds were transferred to a new discrimination (AY+/AX−/CY−). The discrimination involved one unfamiliar compound, AY, which was paired with food. This compound, which comprised a previously relevant signal for food and an irrelevant cue, was expected to elicit a high rate of responding when it was introduced. The discrimination also involved two previously reinforced compounds, AX and CY, that now signaled the absence of food. The associative strengths of AX and CY can be assumed to be similar and high at the end of the preliminary training. If the rate of acquisition of the new discrimination was determined simply by the overall associative strengths of the three compounds (e.g., Pearce, 1987, 1994; Rescorla & Wagner, 1972), then responding to AY could be expected to be sustained at a high level, whereas the rate of responding to AX and CY would gradually decline at the same rate. Alternatively, if the original training resulted in the color-relevant group's attending preferentially to colors A, B, C, and D, and in the pattern-relevant group's attending preferentially to patterns W, X, Y, and Z, there would be differences in the rate at which extinction progresses to AX and CY in the two groups. The discrimination between AY+ and AX− for the color-relevant group is based on two patterns that were previously irrelevant. This discrimination should therefore be hard relative to the pattern-relevant group, for whom patterns were previously relevant. Consequently, extinction to AX should progress more rapidly in the pattern-relevant group than in the color-relevant group. Conversely, the discrimination between AY+ and CY− for the color-relevant group is based on two colors that were previously relevant. This discrimination should therefore be easy relative to the pattern-relevant group, and extinction to CY should progress more rapidly in the color-relevant than in the pattern-relevant group. There were 32 trials with AY and 16 trials with each of AX and CY in each of the nine test sessions of Stage 2.

The results from the first phase of training are presented in Figure 9, which shows that the initial discrimination was more difficult for the pattern-relevant than for the color-relevant group. The results for the two groups
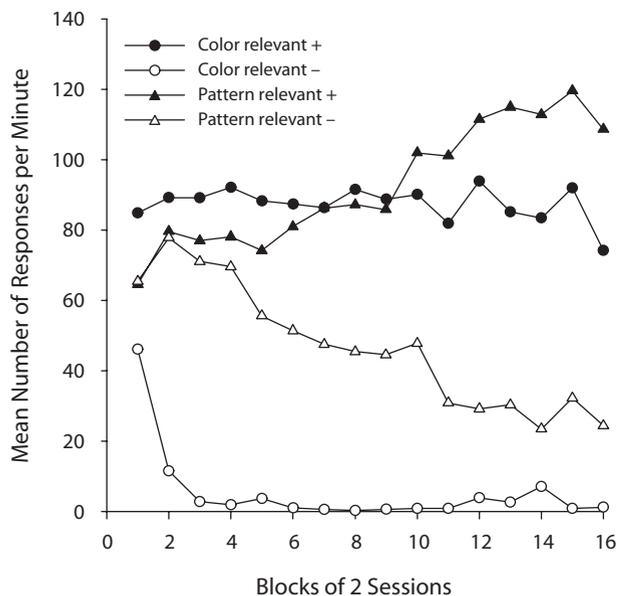
Figure 9. The mean number of responses per minute by the two groups in Experiment 3 during reinforced and nonreinforced trials of the true discrimination in Stage 1.

from the test phase are shown in Figure 10. The rates of responding during the test trials with AY are presented in the left-hand panel, where it is shown that responding was more vigorous by the pattern-relevant than by the color-relevant group. This difference was also evident at the end of Stage 1. The center panel shows the rate of responding to the two compounds for which a pattern was the distinctive cue—that is, the cue that distinguished the compound from AY. It is quite clear from Figure 10 that the rate of responding to AX extinguished more rapidly in the pattern-relevant than in the color-relevant group, which indicates that the discrimination between AY and AX progressed more easily when the distinctive feature had previously

been relevant, rather than irrelevant. In other words, the results imply that one effect of the original training was to encourage pigeons to pay more attention to relevant than to irrelevant stimuli. Moreover, in contrast to the outcome of a number of previous studies, these results cannot be easily explained in terms of acquired distinctiveness. Both X and Y signaled food during Stage 1 for the pattern-relevant group, which should have resulted in their being treated as being equivalent. Such an effect would then be expected to increase the difficulty of discriminating between X and Y in the test phase, whereas the results indicated that this discrimination was relatively easy for the pattern-relevant group.

The remaining panel in Figure 10 shows the results for the test compound CY, in which color was the distinctive cue for the discrimination with AY. The loss of responding by both groups was somewhat faster when color was the distinctive cue than when a pattern was the distinctive cue, but, more important, there is no indication that extinction progressed more readily with the compound containing a previously relevant color rather than that containing a previously irrelevant color. The implication of this result is that although the preliminary training influenced the attention paid to patterns, it had no impact on the attention paid to colors.

We have repeated the previous experiment on several occasions using different colors, different patterns, and with the cues presented in adjacent circles rather than adjacent squares. Even when the colors were quite similar, to our eyes, the outcome was consistently much the same as that portrayed in Figure 10. The results from Experiment 3 are thus reliable and robust. How then, might Spence have explained them? He acknowledged more than once that as they solve a discrimination, animals may learn "preliminary receptor-exposure acts" (Spence, 1952, p. 92), which enable the receptor organs to be directed toward the stimuli that are important for the solution of the discrimination:

In the case of visual discriminations . . . the animal is required to learn, in addition to the final selective
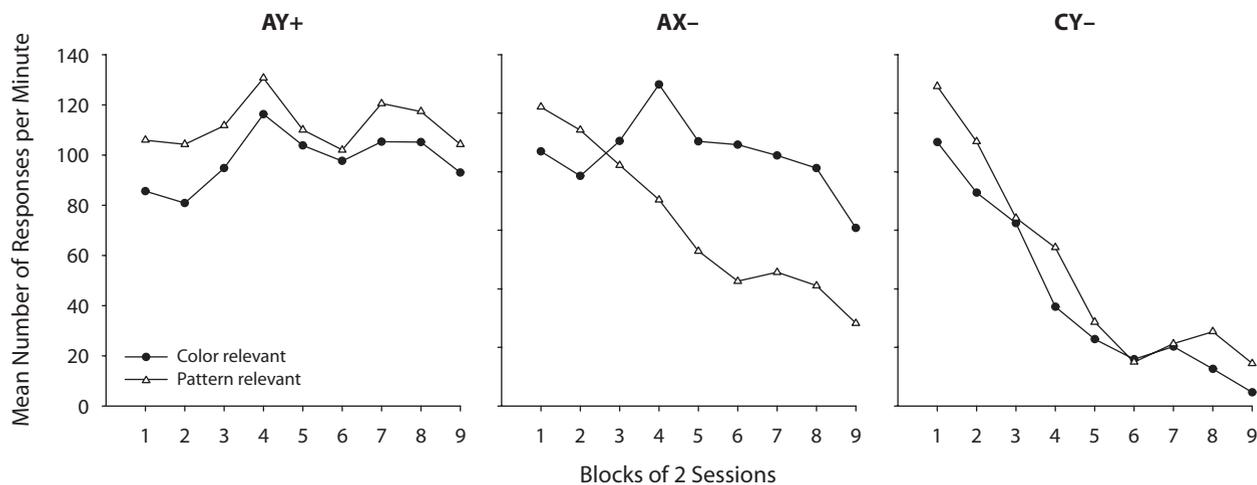

Figure 10. The mean number of responses per minute by the two groups in Stage 2 of Experiment 3 during the reinforced AY trials (left-hand panel), the nonreinforced trials with the compound containing a distinctive cue that was a pattern, AX (center panel), and the nonreinforced trials with the compound containing a distinctive cue that was a color, CY (right-hand panel).

approaching responses, the appropriate (perceptual) response which leads to the reception of the relevant stimulus attributes. That is to say, the animal must learn to orient and fixate its eye so as to receive the critical stimuli. (Spence, 1940, p. 277)

The implication of this proposal for the above experiment is that during the preliminary training, the pigeons learned about the significance of the individual stimuli; in addition, they may also have learned to direct their gaze toward or away from the pattern component of each compound, depending on whether or not it was relevant to the discrimination. As far as the pattern-relevant group is concerned, it appears that unconditioned receptor-exposure acts learned in association with colored squares are so strong that it is not possible to modify them with the method of training that we employed. Alternatively, because the original discriminations were acquired relatively slowly by the pattern-relevant group, more training may have been necessary for that group than for the color-relevant group before the birds could learn to orient toward the relevant cues and away from the irrelevant ones. If this explanation is correct, we would have found a loss of associability to the colors in the pattern-relevant group by giving more training in Stage 1.

Other theorists have argued that the changes in attention to a stimulus are a consequence of more central changes than of the acquisition of orienting responses, as envisaged by Spence (1940, 1952). For example, Sutherland and Mackintosh (1971) assumed the effectiveness of a stimulus to be determined by the extent to which it was detected by an analyzer, which could be switched on or off according to experience. In a related manner, Pearce, George, and Redhead (1998) have outlined a connectionist network for discrimination learning, in which the sensitivity of input units to environmental stimulation is determined by the extent to which they have proven relevant to the solution of a discrimination. The question is then

raised about whether the changes in attention revealed by the previous experiments were brought about by peripheral mechanisms, such as changes in receptor-exposure acts, or by more central mechanisms, such as changes in analyzer strengths. The next experiment was conducted as a preliminary step toward answering this question.

Experiment 4

The stimuli for Experiment 4 were based on those in Experiment 3, except that a single square containing a black pattern superimposed on a colored background was shown on every trial. A drawing of one stimulus can be seen in panel C of Figure 3. The plain colors and patterns were integrated in this way in the hope of making it more difficult for pigeons to look at one component of the compound stimulus while ignoring the other. These stimuli were then used in an exact replication of Experiment 3. The course of acquisition of the discrimination was similar to that shown in Figure 9, even to the extent that responding in the presence of the reinforced compounds was ultimately faster in the pattern-relevant than in the color-relevant group. The results from the reinforced trials with AY during the test phase were also similar to the equivalent results from the previous experiment (see the left-hand panel of Figure 11).

The outcome of the test trials with the two compounds containing a pattern as the distinctive feature was, however, different from that found in Experiment 3. On this occasion, extinction with the compounds occurred equally rapidly in both groups, as the center panel of Figure 11 shows. The figure indicates that responding was slower by the pattern-relevant than by the color-relevant group on some sessions, but this difference fell far from statistical significance. A comparison of these results with those shown in the equivalent panel of Figure 10 indicates that the treatment with the pattern-irrelevant cue disrupted its associability when it was presented in a separate location to the color-relevant cue, but not when it was integrated
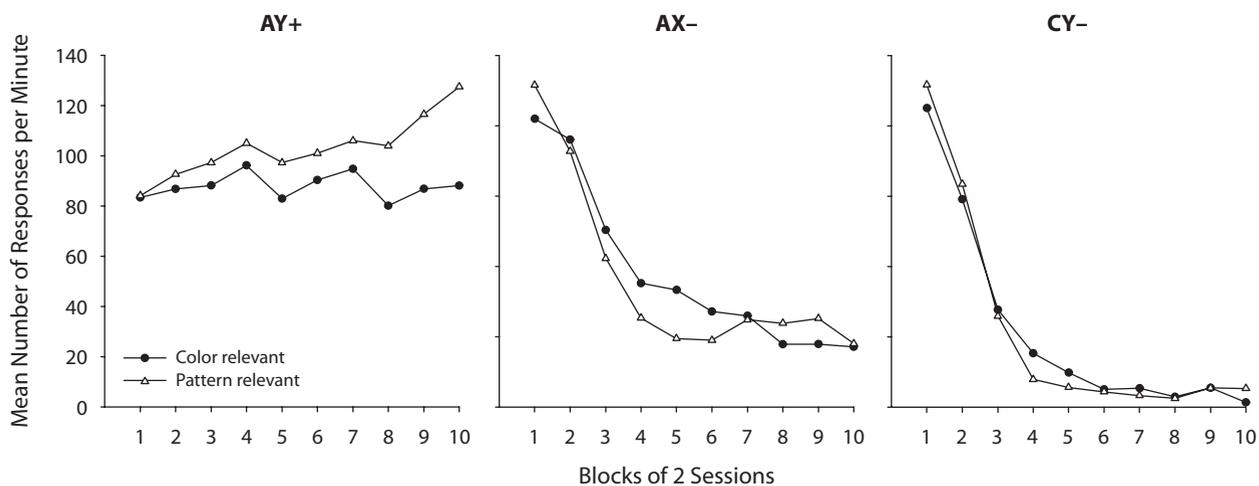

Figure 11. The mean number of responses per minute by the two groups in Stage 2 of Experiment 4 during the reinforced AY trials (left-hand panel), the nonreinforced trials with the compound containing a distinctive cue that was a pattern, AX (center panel), and the nonreinforced trials with the compound containing a distinctive cue that was a color, CY (right-hand panel).

with the color-relevant cue. This result, of course, is entirely congruent with the idea that changes in attention depend on changes in orienting responses.

Presumably, in Experiment 3, pigeons learned to avert their eyes from the patterns when they were irrelevant, but they were unable to learn this strategy in the present experiment because it would also result in their averting their gaze from the relevant color. Also, the results lend no support to the suggestion that changes in attention occur at a more central level than does an orienting response. For these changes to occur, it should not matter whether the components of the compound were presented separately or together, yet the results clearly show that this manipulation has a critically important influence on the outcome of the experiment.

The remaining results from the experiment are presented in the right-hand panel of Figure 11, which shows that the outcome of the test trials with compounds containing a color as the distinctive cue resulted in a similarly rapid loss of responding irrespective of whether the color was previously relevant or irrelevant. Taken together, the results from Experiments 3 and 4 demonstrate that changes in attention occur during discriminations, but only to cues belonging to certain dimensions, and only because pigeons avert their gaze from them.

Experiment 5

The final experiment concerns the relative-validity effect (Wagner et al., 1968), which was described at the beginning of the present article. Although this effect was said to pose a challenge for Spence's (1936, 1937) theorizing, the results from Experiment 4 show one way this effect might be reconciled with his proposal that the learning process is nonselective. We have already seen that this proposal leads to the prediction that the associative strength acquired by X in a true discrimination, $AX+/BX-$, will be the same as in a pseudodiscrimination, $AX\pm/BX\pm$. However, if animals make orienting responses that limit the extent to which they perceive the irrelevant element of a true discrimination, then the capacity for the stimulus to enter into associations may be seriously curtailed. In contrast, with a pseudodiscrimination, in which all stimuli are equally relevant, it is unlikely that animals will selectively ignore the cue that occurs on both trials, and this stimulus will then gain more associative strength than its counterpart in the true discrimination.

To test the above explanation, a true discrimination and a pseudodiscrimination group were trained with stimuli based on those used in Experiment 3. For both groups, A and B were patterns and X was a color presented to pigeons in exactly the same fashion as that described in Experiment 3. The results from Experiment 3 demonstrated that when pigeons are given a true discrimination with a color as the irrelevant cue, they are unable to reduce the attention they pay to it. As a consequence, if the relative-validity effect depends on a loss of attention to the irrelevant cue in the true discrimination, this effect should not be observed in the proposed experiment. The left-hand panel of Figure 12 shows the rates of responding by the two groups during the various trials of the true discrimination and the pseudodiscrimination. Training progressed entirely as would be expected. On completion of this training, both groups received three test trials with color X. As can be seen in the right-hand panel of Figure 12, responding during X was significantly faster in the pseudodiscrimination than in the true discrimination group. This outcome constitutes a demonstration of the relative-validity effect. On the basis of this finding, therefore, it would seem that some mechanism other than a change in attention to the irrelevant stimulus is responsible for the relative-validity effect. A likely cause is that the associative properties of X are not determined solely
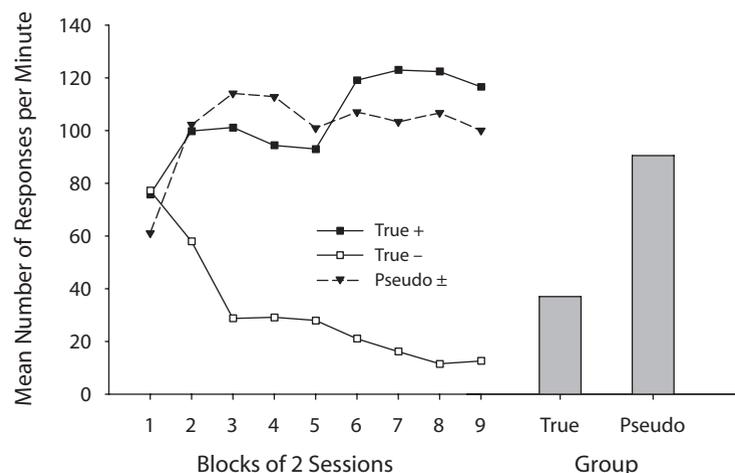

Figure 12. The left-hand side of the figure shows the mean number of responses per minute for the training stage of Experiment 5 during the reinforced AX trials and the nonreinforced BX trials for the true discrimination group, and during the partially reinforced AX and BX trials for the pseudodiscrimination group. The right-hand side of the figure shows the results from the test trials with X alone for the two groups in Experiment 5.

by its relationship with the reinforcer, but also by the associative properties of the stimuli that accompany it (see, e.g., Pearce, 1994; Rescorla & Wagner, 1972). This type of explanation for the relative-validity effect was first advanced by Wagner and Rescorla (1972) over 30 years ago, but to our knowledge, this is the first occasion on which an attentional explanation for the relative-validity effect has been evaluated by using an irrelevant cue that has been shown to be impervious to changes in attention.

Discussion

Experiments 3, 4, and 5 show that the new method we adopted may be helpful for studying how the effectiveness of a stimulus is influenced by discrimination training. Making a stimulus irrelevant in one discrimination appears to reduce the effectiveness of the same stimulus when it is made relevant in a new discrimination. This effect was shown when the stimulus under investigation was a pattern, but not when it was a color. Our results also indicate that a plausible reason for the change in effectiveness of the stimulus is that subjects avert their gaze from the stimulus whenever it is presented. We have to admit to being disappointed by discovering that changes in attention to an irrelevant cue may be mediated by peripheral orienting responses, rather than by more central processes. This disappointment is particularly acute, given our commitment to theories that assume that changes in attention take place after information has entered the nervous system (Pearce et al., 1998; Pearce & Hall, 1980). However, this conclusion is based on the null result shown in the center panel of Figure 11 and should therefore not be taken to imply that changes in attention in pigeons never occur once the perception of a stimulus has taken place. It is worth noting that Mackintosh and Little (1969) were able to demonstrate a superior intradimensional relative to an extradimensional shift with pigeons using stimuli not too dissimilar to those we used in Experiment 4. This result has, however, proven difficult to replicate (Couvillon, Tennant, & Bitterman, 1976; Hall & Channell, 1985), and it remains unclear whether changes in attention in pigeons are possible at a central level.

A word of caution is also needed concerning the claim that it is not possible to change the attention that is paid to color cues when they are irrelevant to the solution of a discrimination. Again, this conclusion has the shortcoming of being based on a null result. Perhaps it would have been possible to modify the associability of colors had the stimuli been presented in a different manner, or had the Stage 1 training been extended.

When Spence (1940, 1942, 1952) proposed that animals make receptor-exposure acts, he assumed that they were strengthened by primary and secondary reinforcers. There may, however, be a problem with applying this assumption to autoshaping. Given an AX+/BX− discrimination, it is reasonable to suppose that food at the end of an AX+ trial will strengthen the receptor-exposure act of looking at A. The question then arises about what would happen on BX trials. Since the act of looking at B would never be followed by food, it is not clear whether attention

to B will be increased. Thus, the analysis of attention that we have offered implies that attention will increase to A, but not B, in an AX+/BX− discrimination. The results from an autoshaping experiment by Wasserman and Anderson (1974; see also Wasserman, 1974) are consistent with this suggestion. Pigeons received an AX+/BX− discrimination in which, for every trial, the two stimuli were presented on different keys, which permitted them to record the rate of responding at A, B, and X separately. As would be expected, responding to A was substantially faster than to B. In addition, during the initial training, responding to X was more vigorous on BX than on AX trials. A ready explanation for this outcome is that the birds oriented toward A whenever it was presented and hardly noticed X on AX trials. In contrast, when B was presented, they may have turned their gaze away from this stimulus, thereby increasing the likelihood of X's being detected, and then pecked on BX trials. The results could also be explained in terms of competing responses: By pecking more frequently at A than at B, the birds would have more time available to peck X on BX than on AX trials. The results from the study by Wasserman and Anderson do not, therefore, provide compelling support for the possibility that an AX+/BX− discrimination results in a stronger orienting response to A than to B. If subsequent experiments should reveal, to the contrary, that this training results in the strength of the orienting response to A being similar to that to B, then explaining how receptor-exposure acts directed at B are strengthened will pose an intriguing theoretical challenge.

No theory is perfect. As far as Spence's theory is concerned, there exist a number of limitations that prevent it from providing a viable account of how animals solve discriminations. A particular shortcoming is its nonselective nature, which means it is unable to explain satisfactorily such effects as blocking (Kamin, 1969), the solution of feature-positive discriminations (Wagner, 1969), and, of course, the relative-validity effect (Wagner et al., 1968). Even so, we have tried to show that at least some of Spence's ideas remain relevant to our current understanding of how animals solve discriminations. The interaction between excitatory and inhibitory generalization gradients shown in Figure 1 provides a useful framework for appreciating how animals solve discriminations between stimulus configurations, and Spence's proposals concerning receptor-exposure acts require further consideration, if our appreciation of attentional processes in animals is to be complete. The study of discrimination learning represents one of psychology's more enduring theoretical endeavors. Spence's theory has already made a significant contribution to this endeavor, and it seems likely that it will continue to do so for many years to come.

REFERENCES

ATKINSON, R. C., & ESTES, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 121-268). New York: Wiley.

COUVILLON, P. A., TENNANT, W. A., & BITTERMAN, M. E. (1976). Intradimensional vs. extradimensional transfer in the discriminative learning of goldfish and pigeons. *Animal Learning & Behavior*, **4**, 197-203.

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

HALL, G., & CHANNELL, S. (1985). A comparison of intradimensional and extradimensional shift learning in pigeons. *Behavioural Processes*, **10**, 285-295.

HULL, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. New York: Appleton-Century.

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, NS: Dalhousie University Press.

KINDER, A., & LACHNIT, H. (2003). Similarity and discrimination in human Pavlovian conditioning. *Psychophysiology*, **40**, 226-234.

KRECHEVSKY, I. (1932). "Hypotheses" in rats. *Psychological Review*, **39**, 516-532.

LASHLEY, K. S. (1929). *Brain mechanisms and intelligence: A quantitative study of injuries to the brain*. Chicago: University of Chicago Press.

LAWRENCE, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology*, **39**, 770-784.

LAWRENCE, D. H. (1950). Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, **40**, 175-188.

MACKINTOSH, N. J. (1974). *The psychology of animal learning*. London: Academic Press.

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

MACKINTOSH, N. J., & LITTLE, L. (1969). Intradimensional and extradimensional shift learning by pigeons. *Psychonomic Science*, **14**, 5-6.

MCLAREN, I. P. L., & MACKINTOSH, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, **28**, 211-246.

MCLAREN, I. P. L., & MACKINTOSH, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, **30**, 177-200.

PEARCE, J. M. (1987). A model for stimulus generalization in 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 & Behavior*, **30**, 73-95.

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: American Psychological Association.

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.

REDHEAD, E. S., & PEARCE, J. M. (1995). Similarity and discrimination learning. *Quarterly Journal of Experimental Psychology*, **48B**, 46-66.

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

SIEGEL, S. (1969). Discrimination overtraining and shift behavior. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 187-213). London: Academic Press.

SPENCE, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, **43**, 427-449.

SPENCE, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, **44**, 430-444.

SPENCE, K. W. (1938). Gradual versus sudden solution of discrimination problems by chimpanzees. *Journal of Comparative Psychology*, **25**, 213-224.

SPENCE, K. W. (1940). Continuous versus non-continuous interpretations of discrimination learning. *Psychological Review*, **47**, 271-288.

SPENCE, K. W. (1942). The basis of solution by chimpanzees of the intermediate size problem. *Journal of Experimental Psychology*, **31**, 257-271.

SPENCE, K. W. (1952). The nature of the response in discrimination learning. *Psychological Review*, **59**, 89-93.

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

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, NS: Dalhousie University Press.

WAGNER, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology*, **56B**, 7-29.

WAGNER, A. R. (2008). Evolution of an elemental theory of Pavlovian conditioning. *Learning & Behavior*, **36**, 253-265.

WAGNER, A. R., LOGAN, F. A., HABERLANDT, K., & PRICE, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, **76**, 171-180.

WAGNER, A. R., & RESCORLA, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301-336). London: Academic Press.

WASSERMAN, E. A. (1974). Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. *Journal of Experimental Psychology*, **103**, 284-297.

WASSERMAN, E. A., & ANDERSON, P. A. (1974). Differential autoshaping to common and distinctive elements of positive and negative discriminative stimuli. *Journal of the Experimental Analysis of Behavior*, **22**, 491-496.

WOODBURY, C. B. (1943). The learning of stimulus patterns by dogs. *Journal of Comparative Psychology*, **35**, 29-40.