THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 1999, *52B* (2), 97–124

Elemental and Configural Processes in Patterning Discrimination Learning

Andrew R. Delamater, Wendy Sosa, and Michael Katz
Brooklyn College–CUNY, New York, U.S.A.

Three experiments used appetitive conditioning with rats to examine the involvement of elemental and configural processes in positive and negative patterning discriminations. The first experiment demonstrated that negative and, to some extent, positive patterning discriminations were learned more rapidly when these discriminations consisted of stimulus elements that had previously been the reinforced as opposed to the non-reinforced elements of a simple discrimination. Experiment 2 revealed an excitatory summation effect during the early phase of negative patterning learning that depended upon discrimination pretraining. The final experiment demonstrated faster discrimination learning between the compound and the less salient, rather than the more salient, element of an instrumental patterning task. The present set of results were interpreted as reflecting the possibility, consistent with connectionist theory, that internal representations of the conditioned stimuli change over the course of a patterning discrimination.

An important problem in conditioning theory continues to involve the specification of how organisms learn about compound stimuli. Two procedures that illustrate different approaches to the learning about compound stimuli are positive and negative patterning. In the former discrimination, a compound stimulus is reinforced while the stimulus elements that compose that compound are non-reinforced when presented in isolation. Conversely, during negative patterning, the compound is non-reinforced while the stimulus elements are separately reinforced.

Three different views about compound conditioning have emerged from the study of patterning discriminations, all of which make different assumptions about how organisms process information contained within a compound stimulus. The “elemental” approach, as exemplified in the research of Rescorla, Wagner, and Kehoe (e.g. Bellingham, Gillette-Bellingham, & Kehoe, 1985; Kehoe & Graham, 1988; Rescorla, 1972, 1973; Rescorla, Grau, & Durlach, 1985; Whitlow & Wagner, 1972) assumes that animals separately represent and learn about each of the stimulus elements that define stimulus compounds.

Requests for reprints should be sent to Andrew R. Delamater, Psychology Department, Brooklyn College–CUNY, 2900 Bedford Ave, Brooklyn, New York 11210 U.S.A. E-mail: andrewd@brooklyn.cunyedu

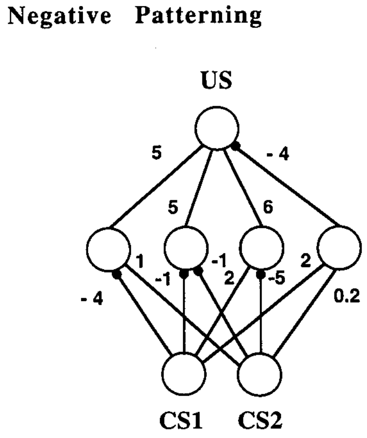
This research was supported through PSC–CUNY grants (# 665527 and # 667519) awarded to the author. The authors would like to thank John Pearce, two anonymous reviewers, and Vin LoLordo for helpful comments on an earlier version of the manuscript.

According to this approach, in order to solve the negative patterning problem organisms must develop excitatory associative strength to each of the stimulus elements, and they must acquire inhibitory associative strength to a compound unique cue that is assumed to be generated by the conjoint presentation of the two conditioned stimuli. The "configural" approach is currently based largely upon the ideas of Pearce (e.g. Pearce, 1987, 1994; Redhead & Pearce, 1995; see also Medin & Schaffer, 1978). According to this view, subjects represent compound stimuli holistically and as being different from but similar to their components. Negative patterning learning amounts to animals acquiring a discrimination between three different, but similar, stimuli. A third approach to the issue of compound learning arises from multi-layered connectionist network models of conditioning (e.g. Gluck & Myers, 1993; Kehoe, 1988; Rudy & Sutherland, 1992; Rumelhart, Hinton, & Williams, 1986; Schmajuk & DiCarlo, 1992). According to this approach, it is assumed that conditioned stimulus representations change during conditioning, and that configural and/or elemental solutions develop according to the nature of the task.

One fairly straightforward manner in which the multi-layered connectionist approach differs from the former approaches is in its assumption about the learning of compound configurations. The elemental and configural views assume that multiple stimuli within a compound spontaneously configure, such that learning about these configural stimuli (be they unique cues or holistic configurations) proceeds from the outset of training. In contrast, the connectionist approach assumes that these "unique cues" or "configural stimuli" are not present from the outset of training but rather are themselves the product of learning.

To illustrate better this aspect of multi-layered connectionist approaches let us consider a fairly typical multi-layered network solution to the negative patterning problem. Figure 1 depicts the model network solution that is produced when a 2:4:1 network (input units: hidden units: output unit) is trained using a standard back-propagation training algorithm (see Rumelhart et al., 1986). Computational details regarding this simulation can be found in Appendix A.

Notice that the conditioned stimuli, CS1 and CS2, are represented by input units, that the unconditioned stimulus (US) is represented as a single output unit, and that internal representations of the conditioned stimuli are conceived of as the pattern of activation of hidden layer units on a given trial. The connection values between various units, which emerged in a simulation of a negative patterning task with a stringent learning criterion, are also indicated within the network. Depicted on the lower portion of the figure is the pattern of activation of the four hidden units and the US unit after the network learns the task for separate CS1+, CS2+, and CS1CS2− trials. Notice that in this solution (and this is not an atypical solution to this problem with this network) the pattern of hidden unit activations is different on each trial type. Note further that on compound trials the strongly activated hidden unit itself strongly inhibits the US unit, whereas on single element trials the pattern of activation of hidden layer units results in net excitation of the unconditioned stimulus unit. These patterns of activation across the hidden layer are not present from the outset of training, but they emerge as "internal representations" of the stimuli. Psychologically speaking, when the network produces a different pattern of hidden unit activations on compound and element trials then the network is said to have learned a configural representation of the compound. This view of stimulus configuration



	H1	H2	H3	H4	US
CS1 :	[-	-	+	+]	+
CS2 :	[+	-	-	+]	+
CS1+CS2 :	[-	-	-	+]	-

FIG. 1. Representative 2:4:1 network solution to a negative patterning task using a back-propagation learning algorithm. The lower portion indicates the hidden unit activation pattern and the activation of the unconditional stimulus (US) unit on each trial type.

contrasts sharply, however, with the manner in which other models interpret configural processes (e.g. Pearce, 1994). Configural representations in the standard back-propagation model must be learned.

The present set of studies were concerned with investigating the basic claim of multi-layered connectionist models that internal representations of stimuli change over the conditioning phase. The first experiment examined a set of predictions regarding positive and negative patterning learning that stem uniquely from the multi-layered back-propogation network portrayed in Figure 1 (Rumelhart et al., 1986). The next two experiments further explored the proposal that configural representations are learned by determining if elemental and configural processes might contribute to negative patterning learning at different times during acquisition.

EXPERIMENT 1

A potentially useful manipulation is suggested by the network approach described above to evaluate the suggestions that configural representations must be learned in solving patterning discriminations. A pretraining manipulation that partially generates internal representations of the stimuli that resemble those required for solving the patterning problem should facilitate patterning learning.

One pretraining manipulation that was investigated in the present study involved a comparison between the effects of reinforced and non-reinforced pretraining upon the learning of positive and negative patterning discriminations. Consider a situation in which subjects are taught to make an auditory discrimination (A) and a visual discrimination (V) in a pretraining phase (e.g. A1+, A2−, V1+, V2−). During a subsequent phase, different groups could be trained on a negative patterning task involving either the previously reinforced (A1+, V1+, A1V1−) or the previously non-reinforced stimuli (A2+, V2+, A2V2−). Similarly, different groups of subjects could be trained on a positive patterning task after reinforced (A1−, V1−, A1V1+) or non-reinforced (A2−, V2−, A2V2+) pretraining.

The multi-layered network model depicted in Figure 1 predicts that the patterning discriminations will proceed rapidly if the consequence of the compound during patterning training was also the consequence for its components in the pretraining phase. In other words, this model predicts that reinforced pretraining will facilitate positive patterning learning, and that non-reinforced pretraining will facilitate negative patterning learning. Simulation results are presented in the top panel of Figure 2. Details regarding this simulation can be found in Appendix B.

Fortunately, the reason for these predictions is not obscure. During the pretraining phase, associations between input and hidden unit layers develop such that internal representations of each of the similarly treated stimuli become “compressed” at the hidden layer (see also Gluck & Myers, 1993). This means that each stimulus that is treated alike (e.g. A1+ and V1+ or A2− and V2−) comes to activate the same set of hidden units during the pretraining phase. These hidden units, in turn, develop the appropriate associations with the US unit (activating it on A1+ and V1+ trials, and inhibiting it on A2− and V2− trials). Thus, by the time patterning training begins, the compound stimulus already strongly evokes an internal representation that is appropriately connected to the US unit in solving the patterning task. The network requires relatively few additional weight adjustments in order to solve the patterning problem. The simulation results clearly show that non-reinforced pretraining facilitates negative patterning relative to reinforced pretraining, and that reinforced pretraining facilitates positive patterning learning.

In the present experiment two groups of rats were trained initially to discriminate between two auditory stimuli and two visual stimuli in the same conditioning sessions (A1+, A2−, V1+, V2−). Different sets of subjects were then trained with a positive patterning task in Experiment 1a and with a negative patterning task in Experiment 1b. In each of these experiments, group reinforced pretraining was trained in the patterning task using the stimuli that had been reinforced earlier in the initial simple discrimination phase. Similarly, in each of these experiments, group non-reinforced pretraining was trained in the patterning task using stimuli that had been non-reinforced in the earlier phase.

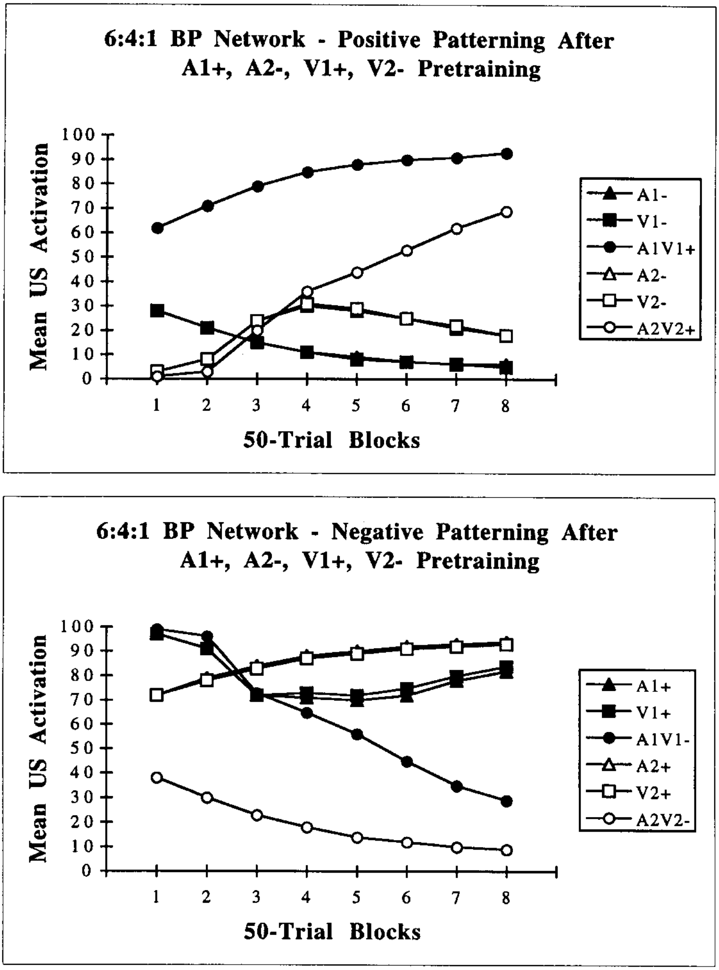


FIG. 2. Simulation results for negative patterning (lower panel) and positive patterning (upper panel) tasks solved by a 6:4:1 network model. See text for more details.

Experiment 1a: Effects of Pretraining upon Positive Patterning Learning

Method

Subjects

Sixteen male Sprague-Dawley rats (supplied by Charles River Breeders) weighing approximately 340 g at the beginning of the experiment were used. They were individually housed in a colony room that was on a 16-hr light/ 8-hr dark cycle, and they were maintained with daily supplemental feedings at 85% of their ad lib body weights.

Apparatus

The apparatus consisted of eight identical standard conditioning chambers, each of which was housed in a sound- and light-resistant shell. The conditioning chambers measured 30.5 cm × 24.0 cm × 25.0 cm. Two end walls were constructed of aluminium, and the side walls and the ceiling were made from clear Plexiglas. The floor consisted of 0.60-cm diameter stainless steel rods spaced 2.0 cm apart. In the centre of one end wall 1.2 cm above the grid floor was a recessed food magazine measuring 3.0 × 3.6 × 2.0 cm (length × width × depth). A single 45-mg pellet (P.J. Noyes Co., Formula A) was dropped onto the magazine floor when the reinforcer was scheduled. On the inner walls of the recessed magazine were an infra-red detector and emitter enabling the automatic recording of head movements inside the magazine. These were located 0.9 cm above the floor and 0.8 cm recessed from the front wall. Located 3.0 cm to the right of the magazine and 8.0 cm above the floor was a lever. Access to the lever was prevented in these experiments by a sheet-metal covering. A 6-W light bulb was mounted near the top of the rear wall of the outer chamber, above and towards the rear end of the conditioning chamber. When activated, this light-bulb flashed, with equal on-off periods, at approximately 1.5 cycles/sec. A second 6-W light-bulb was mounted on the bottom of the side wall of the outer chamber, below and behind the rear wall of the conditioning chamber. When activated, this light-bulb remained continuously illuminated. Two speakers were mounted approximately 22 cm behind the front wall of the conditioning chamber (where the food magazine was located). A 1500-Hz pure tone generated by the computer and amplified by a Radio Shack amplifier was presented through one of these speakers. A white noise stimulus (produced by a Grason-Stadler white-noise generator) was presented through the other speaker. These auditory stimuli measured 4 dB (1500-Hz tone) and 12 dB (white noise) above a background level of 78 dB (C weighting). The chamber was dark except when the visual stimuli were presented. Fans attached to the outer shells provided for cross-ventilation within the shell as well as background noise. All experimental events were controlled and recorded automatically by a 386 IBM-clone microcomputer and interfacing equipment (Alpha Products) located in the same room.

Procedure

Prior to the beginning of this experiment the rats were magazine trained with pellet and sucrose reinforcers. In each magazine training session 20 pellet reinforcers were delivered according to a variable time 60-sec schedule. Sucrose reinforcers (0.2 ml of a 16% solution) were delivered to the same food magazine as the pellets, but in a separate session. The sucrose solution was not used in the present study.

Pretraining Over the next 15 sessions, all rats received training on an auditory discrimination and a visual discrimination. In each daily session, two auditory and two visual stimuli were presented on separate trials, but only one of each of these pairs was reinforced with food pellets. For subjects in group reinforced pretraining ($n = 8$), the flashing light (F) and the white noise (N) stimuli were reinforced while the steady light (L) and tone (T) stimuli were non-reinforced. For subjects in group non-reinforced pretraining ($n = 8$), F and N were non-reinforced while L and T were reinforced during these sessions. Each session contained eight trials of each type, and the intertrial interval (ITI) was variable around a mean of 60 sec, ranging from 30 to 90 sec. The stimuli were 10 sec in duration and the delivery of a single 45-mg food pellet was coincident with the offset of reinforced stimuli.

Positive Patterning During the next 16 sessions, all subjects were trained on a positive patterning discrimination task. Each group received trials in which F and N were presented individually as well as in a simultaneous compound (FN). These stimuli were 10 sec in duration as in the pretraining phase, and a single pellet was delivered coincident with the offset of the compound stimulus. When presented individually, F and N were non-reinforced. There were eight trials of each type in each daily session, and these were irregularly interspersed throughout the session with the constraint that each trial type occurred in each of eight 3-trial blocks. Four different trial sequences were used throughout the experiment. The ITI averaged 2 min and ranged between 1 and 3 min.

Statistical Analysis. Analysis of variance (ANOVA) techniques were used here and throughout the paper to evaluate the data. In order to gain added statistical power in detecting possible interactions, the *F* tables of Rodger (1974, 1975) were used in evaluating main effects and interactions.

Results

Magazine response rates were recorded both during the stimuli and during a 10-sec prestimulus period here and throughout all the experiments reported. Acquisition of the Phase 1 discriminations proceeded uneventfully. By the end of Phase 1, all subjects responded more to the reinforced than to the non-reinforced stimuli. Over the final three sessions of Phase 2, group reinforced pretraining averaged 32.4 responses per min during the reinforced stimuli and 5.9 responses per min during the non-reinforced stimuli. Group non-reinforced pretraining averaged 33.7 and 7.7 responses per min, respectively, during reinforced and non-reinforced trials.

The course of the positive patterning discrimination for both groups is displayed in Figure 3. Mean magazine responding during FN, F, and N is displayed over successive two-sessions blocks. It is apparent from this figure that the two groups acquired the discrimination at similar rates. A Group × Stimulus × Block ANOVA revealed a significant stimulus main effect, $F(2, 28) = 33.21$, as well as a significant Stimulus × Block interaction, $F(14, 196) = 8.78$. These results suggest that the discrimination improved with training. The significant Group × Block interaction, $F(7, 98) = 3.26$, reflects higher overall responding on the first block in the group given reinforced pretraining. The Group × Stimulus × Block interaction was also significant, $F(14, 96) = 1.46$.

A Group × Block ANOVA was also performed on the prestimulus magazine response rate data. This analysis revealed significant main effects for the group, $F(1, 14) = 5.04$ and block, $F(7, 98) = 1.70$, variables. The overall rates of magazine responding during the prestimulus intervals in the two groups were quite low, but more responding occurred in group non-reinforced pretraining (3.0 responses per min) than in group reinforced pretraining (1.8 responses per min).

Close inspection of the positive patterning discrimination data revealed that a single subject in group nonreinforced pretraining responded at an extremely high rate to the reinforced compound. Averaging over the final four 2-session blocks, this subject responded at a rate more than two standard deviations higher than all the other subjects in that group. This subject's mean response rate was 89.5 responses per min compared to the group mean of 34.3 responses per min. In order to evaluate more fully the source of the three-way interaction mentioned above and to minimize the contribution of this single deviant subject, the data were transformed to discrimination ratios. As there was no difference in responding to the elements, the discrimination ratio was of the form A/

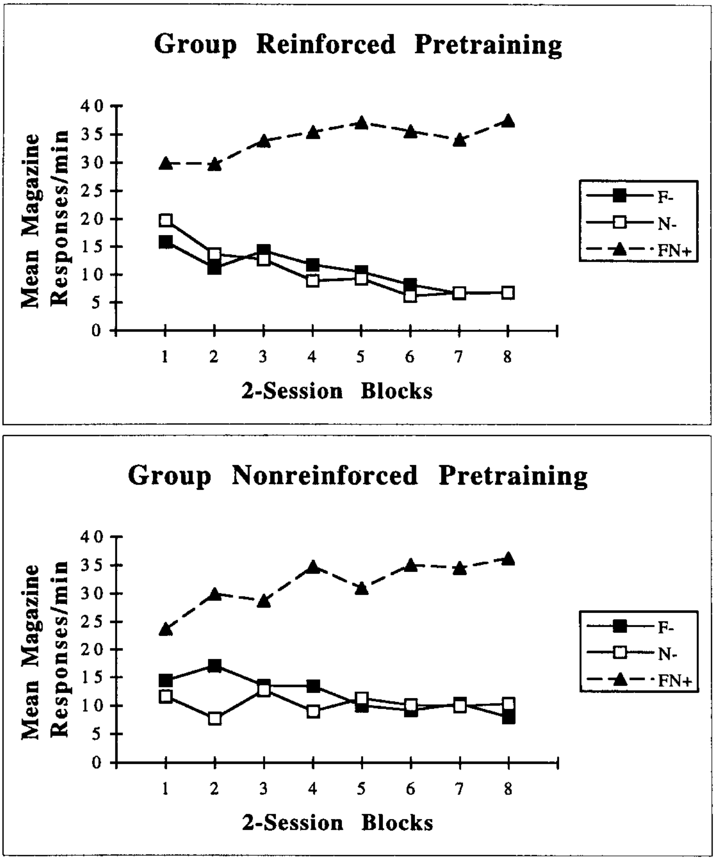


FIG. 3. Mean rates of responding across two-session blocks in Experiment 1a to each of the stimuli in the positive patterning task following reinforced (upper panel) or non-reinforced (lower panel) pretraining to the elements in the prior simple discrimination phase.

(A+B), where A refers to responding to the compound, and B refers to responding averaged over the two elements. These transformed data are portrayed for each group over successive blocks of training in Figure 4.

The discrimination ratio data indicate more clearly that the two groups acquired the positive patterning discrimination at different rates. This impression was confirmed with an ANOVA, which revealed a significant Group × Block interaction, *F*(7, 98) = 1.45. Subsequent between-group ANOVAs were performed on these data at each block using a pooled error variance (Kirk, 1968). This analysis revealed group differences at Blocks 5 and 7, *F*s(1, 31) = 4.92. The apparent differences on Blocks 6 and 8 just fell short of statistical significance, *F*s(1, 31) = 3.95 and 3.28, respectively, *P*s < .10.

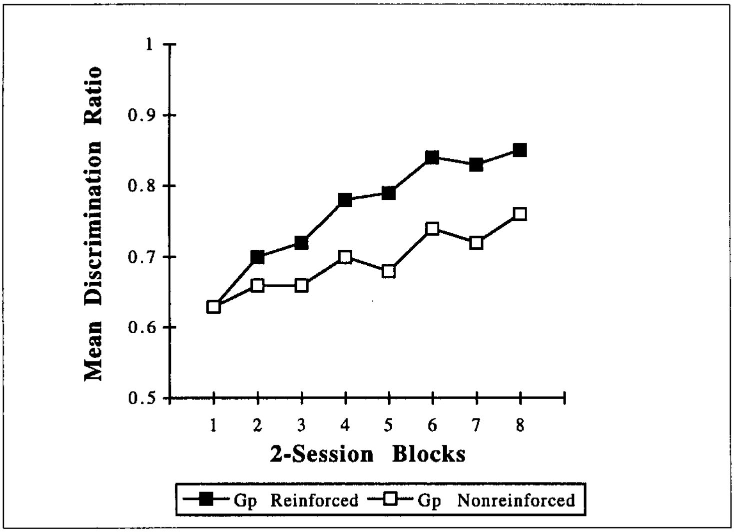


FIG. 4. Mean discrimination ratios for the groups of Experiment 1a solving the positive patterning task after having received reinforced or non-reinforced pretraining to the elements.

Discussion

The present data offer modest support for the prediction derived from the Rumelhart et al. (1986) network model that faster learning of the positive patterning task should arise from reinforced pretraining to the elements. Although this difference was not obvious when considering the response rates, transformation of these data to discrimination ratios did reveal the predicted group difference. To evaluate more fully the claims of that model, however, it is necessary to examine the effects of reinforced and non-reinforced pretraining upon negative patterning learning. Experiment 1b explored that issue.

Experiment 1b: Effects of Pretraining upon Negative Patterning Learning

Method

Subjects

Sixteen new male Sprague-Dawley rats (supplied by Charles River Breeders) weighing approximately 350 g at the beginning of the experiment were used. They were housed and maintained as in Experiment 1a.

Apparatus

The apparatus was the same as that used in Experiment 1a.

Procedure

Magazine training was the same as that used for Experiment 1.

Pretraining Pretraining was carried out exactly as it had been in Experiment 1a.

Negative Patterning Training During the next 16 sessions, all subjects were trained on a negative patterning discrimination using the F and N stimuli from the pretraining phase. The procedure was similar in most respects to the positive patterning procedure used in Experiment 1a with a few noted exceptions. In this task reinforcement occurred after the stimuli when they were presented in isolation but not when they were presented in compound. Each session contained 4 reinforced trials with each of F and N and 16 non-reinforced trials with FN, with a mean ITI of 2 min. These trials were irregularly interspersed throughout the session, with the constraint that each of four 6-trial blocks included 4 non-reinforced FN trials, 1 reinforced F trial, and 1 reinforced N trial. Four different trial sequences were used in the experiment. These parameters were chosen in order to match the positive and negative patterning tasks on the total number of trials and reinforced trials in each session.

Results

Acquisition of the Phase 1 discriminations in Experiment 1b proceeded smoothly. By the end of Phase 1 all subjects responded more to the reinforced than to the non-reinforced stimuli. Over the final three sessions of Phase 1, group reinforced pretraining averaged 25.1 responses per min during the reinforced stimuli and 6.8 responses per min during the non-reinforced stimuli. Similarly, group non-reinforced pretraining averaged 34.6 and 8.2 responses per min, respectively, during reinforced and non-reinforced trials. The apparent difference between the groups in responding on reinforced trials was not statistically reliable, $F(1, 14) = 2.96, p > .10$.

Performance on the negative patterning discrimination is shown for each group in Figure 5. Presented is the mean rate of magazine responding during F, N, and FN over successive two-session blocks of training. There are three aspects of these data worth emphasizing. First, group reinforced pretraining (top panel) appears to have learned the discrimination more rapidly than group non-reinforced pretraining (bottom panel). Successful discrimination between each element and the compound occurred sooner in group reinforced pretraining. Second, in both groups it was true that discrimination learning was more rapid between the auditory element and the compound than between the visual element and the compound. Third, an excitatory summation effect occurred early in training. During the first block there was more responding during the compound than during either element. This effect is more apparent in group non-reinforced pretraining.

Many of these impressions were confirmed with a Group × Stimulus × Block ANOVA performed on these data. Significant stimulus, $F(2, 28) = 3.05$, and Stimulus × Block, $F(16, 224) = 13.24$, effects revealed that over the course of training subjects learned to discriminate among the three stimuli. A significant Group × Block interaction, $F(8, 112) = 6.77$, indicated that the overall level of responding was greater early in training in group reinforced pretraining. Importantly, the significant Group × Stimulus

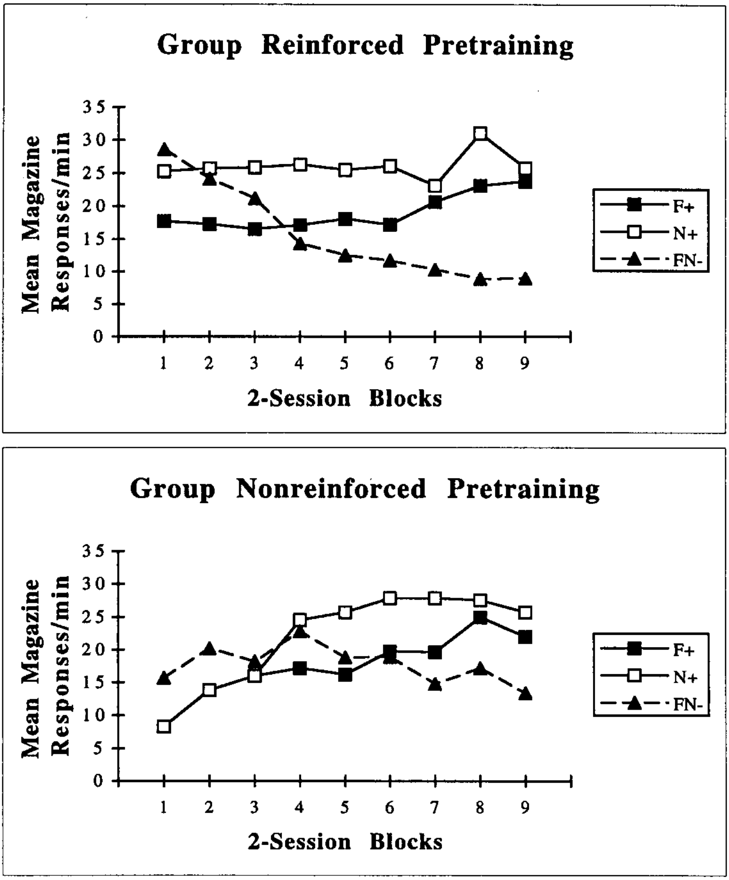


FIG. 5. Mean rates of responding across two-session blocks of Experiment 3 to each of the stimuli in the negative patterning task following reinforced (upper panel) or non-reinforced (lower panel) pretraining to the elements in the prior simple discrimination phase.

× Block interaction, *F*(16, 224) = 1.70, revealed that the groups acquired the discrimination at different rates.

In order to assess the nature of this three-way interaction, separate repeated one-way ANOVAs were performed at each block for each group using a pooled error term (Kirk, 1968). For the group given reinforced pretraining, these tests revealed significant differences in responding on the three types of trials on Blocks 4 through 9, *F*s(2, 35) > 3.14. Subsequent post hoc tests following the methods of Rodger indicated that responding during N exceeded responding during FN in each of these blocks, but that responding during F was greater than responding during FN only in Blocks 7 through 9. Moreover, responding was shown to be greater during N than during F on Blocks 4, 5, 6, and 8. For group non-reinforced pretraining, one-way ANOVAs performed at each block indicated differences in responding on the three types of trials only on Blocks 7 and 9, *F*(2, 35) = 3.53 and 3.23, respectively. Additional post hoc tests revealed more responding during N than during FN on Blocks 7 and 9, and that responding during F exceeded responding during FN only on Block 9.

The excitatory summation effect in the first block did not reach statistical significance in the analysis just reported because of a considerable amount of variability in the data beyond the first block of training. The error variance on the first block, *MS*e = 24.13, was drastically less than the error variance averaged over all subsequent blocks, *MS*e = 139.6. This state of affairs justifies a closer look at performance during this block.

A separate Group × Stimulus ANOVA was performed just on the first block data. This analysis revealed a significant summation effect. For this analysis, responding to the compound for each animal was compared to responding to the element that elicited the higher rate of magazine responding. Thus, greater compound than element responding with this measure would conservatively indicate excitatory summation. In group reinforced pretraining, compound responding was 28.6 responses per min compared to 26.6 responses per min for the element responded to most. In group non-reinforced pretraining, the comparable response rates were 15.7 and 11.7 responses per min for compound and element trials, respectively. The analysis revealed a significant main effect of group, $F(1, 14) = 15.25$, and of stimulus, $F(1, 14) = 5.87$. However, the Group × Stimulus interaction did not approach significance, $F(1, 14) = 0.70$.

A separate Group × Block ANOVA was performed on the prestimulus magazine response rates during negative patterning discrimination training. This analysis revealed a significant Group × Block interaction, $F(8, 112) = 2.26$. This interaction indicates that, as was found in Experiment 1a, somewhat more prestimulus responding occurred in group non-reinforced pretraining (5.4 responses per min) than in group reinforced pretraining (3.9 responses per min), but that this was true for some but not all of the training blocks.

Discussion

The results from Experiments 1a and 1b can be summarized as revealing that reinforced pretraining facilities, negative patterning learning (compared to non-reinforced pretraining), and, with the discrimination ratio measure at least, it also facilitates positive patterning learning. These experiments were motivated by the unique prediction derived from the multi-layered network model described above, which suggests that reinforced pretraining would facilitate positive patterning and that non-reinforced pretraining would facilitate negative patterning. Only the first of these expectations received some support.

The central claim of the network model that gave rise to these predictions is that configural representations are not normally present from the outset of training, but that they are acquired over the conditioning phase. The pretraining manipulation used here was expected to construct partially the configural representations that could then be used in appropriate ways during the positive and negative patterning tasks. The present data suggest that if changes in the internal representations of stimuli occur throughout training, they do not do so in the manner anticipated by the standard multi-layered network model of Rumelhart et al. (1986).

An alternative way of thinking about the present data is to think of pretraining as influencing not the manner in which stimuli are represented but the amount of processing the stimuli receive. For example, suppose that reinforced stimuli are processed more effectively than non-reinforced stimuli. One consequence of this would be to enhance the salience of the stimuli, and this could, in turn, result in faster learning.

Such an idea is not unprecedented (e.g. Mackintosh, 1975). More recently, Rescorla (1991) presented data from a task related to those used here to make a similar point. In Rescorla's study, the modulatory influence that either an inhibitor or a facilitator exerted upon responding to a target stimulus was enhanced by separately reinforcing the modulatory stimulus. These results are especially relevant to the present studies because patterning discrimination learning might itself be understood in terms of the stimuli acquiring modulatory functions. For example, as the negative patterning task consists of A+, B+, and AB− trials, one can construe learning in terms of A inhibiting B (or vice versa) during AB compound trials. Rescorla demonstrated (using a within-subjects design) that in a B+, AB− task separate A+ training facilitated learning relative to separate A− training. A natural interpretation of this result together with those reported here is that separate reinforced training enhanced the processing of A.

Although this interpretation makes sense of the fact that reinforced pretraining facilitates both positive and negative patterning learning, it may be premature to dismiss the central notion of the network approach that internal representations undergo changes across conditioning. Two additional aspects of the results from Experiment 1b are worthy of further exploration.

First, recall that excitatory summation was observed early in training in Experiment 1b. This result is readily explained by an elemental model of conditioning but not by a configural model (e.g. see Redhead & Pearce, 1995). The Rescorla–Wagner model (Pescorla & Wagner, 1972), for example, assumes that compound responding reflects the sum of the associative tendencies conditioned to the elements. In contrast, in Pearce's configural theory, the elements generalize only a portion of their associative strength to the compound. Unless one also assumes that the context contributes to this generalization (e.g. Aydin & Pearce, 1995, 1997; Pearce, Adam, Wilson, & Darby, 1992; Rescorla, 1997), then the total associative strength generalizing to the compound will only equal that conditioned to either stimulus element, and no summation would be obtained. Parenthetically, although the inclusion of context can be used by Pearce's model to generate a summation effect in some procedures, we have been unable to observe such an effect when simulating the effects of non-reinforced pretraining of the sort used here upon negative patterning learning.

Given the theoretical significance of summation during negative patterning, it becomes important to understand the conditions under which one might expect to observe summation in this task. Data from pilot studies conducted in our laboratory have indicated that summation does not occur when negative patterning is not preceded by a simple discrimination. Redhead and Pearce (1995) similarly failed to observe excitatory summation early in negative patterning training using a procedure that did not include pretraining to the elements. Thus, in addition to altering the processing of stimuli, the pretraining manipulation used in Experiments 1a and 1b might somehow alter the manner in which the stimuli are represented, which would encourage excitatory summation.

The second finding from Experiment 1b that remains to be explained is that the discrimination between N+ and FN− was learned more quickly than the discrimination between F+ and FN−. Redhead and Pearce (1995) demonstrated how configural and elemental theories differ in their predictions regarding the relative speed of discriminations between the non-reinforced compound and each of the stimuli when they differ in salience. It becomes of interest, therefore, to ask whether the F and N stimuli used here differed in salience, and if so whether the pattern of results with these stimuli follows that predicted by an elemental or a configural model. Each of the next two experiments addressed these issues in turn.

EXPERIMENT 2

The present experiment determined whether the summation seen at the onset of the negative patterning discrimination depended upon the pretraining procedure used in Experiment 1. In each of Experiments 2a and 2b one group—group non-reinforced pretraining—was given the same training as the similarly named group in Experiment 1b. In Experiment 2a, this group was compared to a control group that received an equal number of exposures to the conditioning context but without stimuli being presented. This control group would determine if summation would occur without pretraining to the stimuli. In Experiment 2b group non-reinforced pretraining was compared to a control group that received non-differentially reinforced pretraining. This control group was exposed to each of the auditory and visual stimuli in each pretraining session, but these stimuli were all partially reinforced (A1+, A1−, A2+, A2−, V1+, V1−, V2+, V2−). This control group would determine whether preexposure to the stimuli and reinforcement, and not the discrimination pretraining per se, would be sufficient for observing summation.

Experiments 2a and 2b

Method

Subjects

32 experimentally naive male Sprague-Dawley rats (supplied by Charles River Breeders) weighing approximately 375 g at the beginning of the experiment were used. Of these, 16 were used in Experiment 2a and 16 in Experiment 2b. The rats were housed and maintained as in Experiment 1.

Apparatus

The apparatus was the same as that used in Experiment 1.

Procedure

Experiment 2b was performed after Experiment 2a had been completed, but the same general procedures were used in both except where noted below. The rats were initially given two magazine training sessions with pellet reinforcement. Each of these sessions was 20 min long, during which 20 reinforcers were delivered randomly in time.

Pretraining For group non-reinforced pretraining ($n = 8$ in Experiment 2a and $n = 8$ in Experiment 2b), the pretraining phase was much the same as that in Experiment 1. Over 16 sessions of pretraining, these subjects received non-reinforced trials with F and N and reinforced trials with L and T. Subjects in the control group of Experiment 2a—group context preexposure—were merely placed in the conditioning chambers for comparable periods of time without any stimulus events occurring during most of this phase. On the 16th day of pretraining, group context preexposure received the same number and temporal distribution of pellet deliveries as those given to group non-reinforced pretraining on that session, in order to reacquaint this group with pellets in the magazine.

The control group of Experiment 2b—group non-differential reinforcement—received each stimulus (N, T, F, and L) reinforced on two of eight trials. In addition, in order to equalize the total number of reinforcements received in the two groups, there were eight presentations of the reinforcer in the ITI. These unsignalled reinforcers were presented during the ITI with the constraint that they be separated from the immediately preceding and following trials by between 30 and 65 sec. With these parameters, the average time separating successive reinforcements (“cycle time” = 120 sec) is three times greater than the average time to reinforcement within the stimuli (“trial time” = 40 sec). Thus, it was expected that by the end of pretraining the two groups would be responding comparably and at low levels to F and N.

Negative Patterning During the next four sessions, all subjects were trained on a negative patterning discrimination task using the F and N stimuli from the pretraining phase. The procedures were identical to those used in Experiment 1b. Briefly, each session contained 4 reinforced F trials, 4 reinforced N trials, and 16 non-reinforced FN trials, with a mean ITI of 2 min.

Results

Acquisition of the Phase 1 discriminations proceeded smoothly. By the end of Phase 1 training, group non-reinforced pretraining responded more to the reinforced than to the non-reinforced stimuli. In Experiment 2a these subjects averaged 33.1 and 6.7 response per min, respectively, during reinforced and non-reinforced trials over the final three sessions of Phase 1. In Experiment 2b group non-reinforced pretraining averaged 29.9 and 7.7 response per min, respectively, during reinforced and non-reinforced trials over the final three sessions of Phase 1. Group non-differential reinforcement (from Experiment 2b) responded on average 6.9 responses per min to the stimuli (flash and noise) to be used during negative patterning training.

Performance averaged across the four sessions of the negative patterning discrimination is shown for each group in Figure 6. The mean rate of magazine responding during F, N, and FN is shown separately for the groups run in Experiment 2a (top panel) and Experiment 2b (bottom panel). As was true in Experiment 1b, the two groups given non-reinforced pretraining once again displayed evidence of summation—that is, they responded more during the compound than during either element. In contrast, the two control groups responded either less or no differently to the compound than to the element responded to most (the noise).

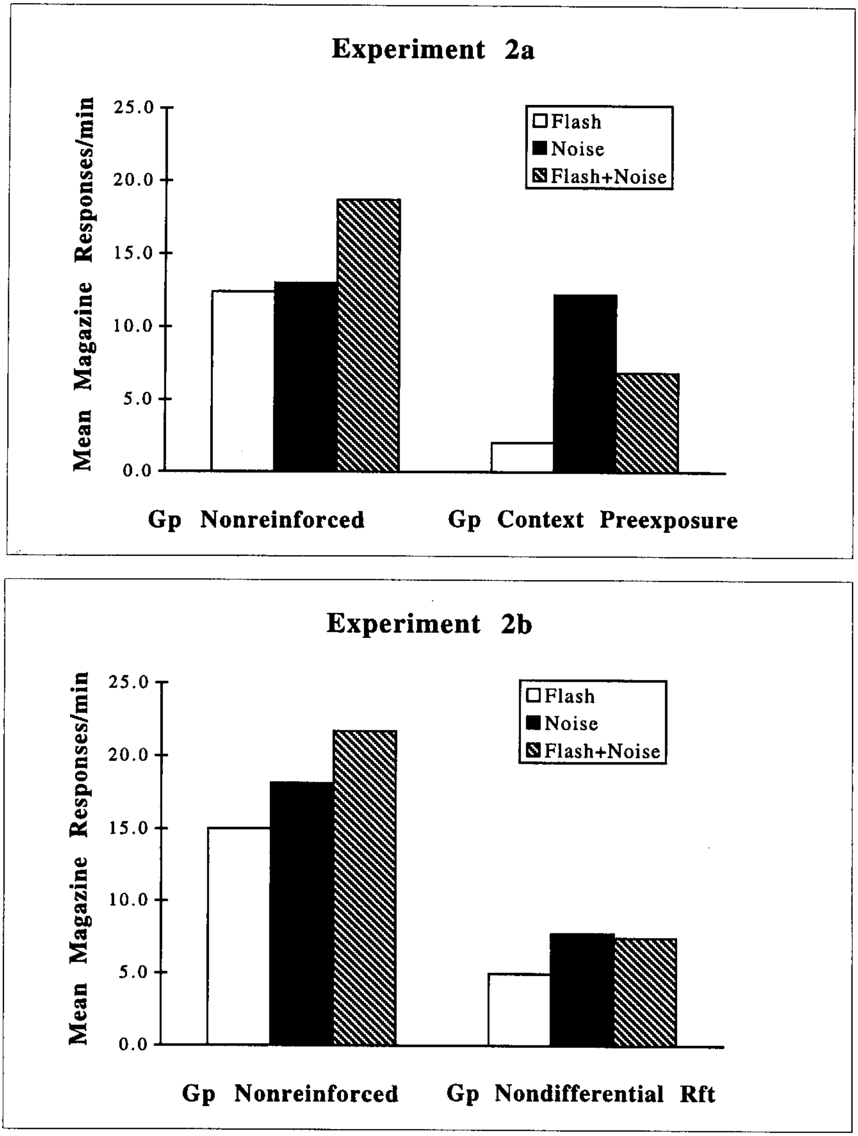


FIG. 6. Mean rates of responding to each stimulus during the four sessions of negative patterning training for group non-reinforced pretraining and for the control groups of Experiment 2a (upper panel) and Experiment 2b (lower panel).

These data were analysed with a Group × Stimulus ANOVA. For this analysis, data from the two groups given non-reinforced pretraining were combined to form one level of the group factor, and the two control groups were included as additional levels of this factor. For the purposes of evaluating summation, the measure reported in Experiment 1b was adopted here. In particular, responding to the compound stimulus for each subject was compared to responding to the stimulus element that controlled the higher rate of responding. This analysis revealed a significant main effect of group, $F(2, 29) = 4.83$, as well as a significant Group × Stimulus interaction, $F(2, 29) = 11.89$. The group main effect indicates that there was more overall responding in group non-reinforced pretraining than in the control groups. The interaction was evaluated with separate repeated measures analyses performed on each group using a pooled error term. This analysis revealed that compound responding exceeded element responding in group non-reinforced pretraining, $F(1, 29) = 9.58$, that it was less than element responding in group context preexposure, $F(1, 29) = 13.66$, and that it did not differ from the element that was most responded to in group non-differential reinforcement, $F < 1$.

A separate between-group ANOVA was performed on the prestimulus magazine response rate data. Although prestimulus responding was numerically greater in group non-reinforced pretraining (4.5 responses per min) than in group context preexposure (3.4 responses per min) and group non-differential reinforcement (2.6 responses per min), this analysis failed to reveal any between-group differences.

Discussion

The results of Experiments 2a and 2b support the finding that discrimination pretraining results in summation during the early stages of negative patterning learning. Moreover, this enhanced responding to the compound cannot be attributed to an effect of mere exposure to the stimuli or to the context during the pretraining phase. It seems likely that this summation effect in the negative patterning procedure depended upon the stimuli having undergone prior discrimination training. However, it may be noted that firm conclusions regarding the control groups of the present experiment are complicated by the fact that these groups responded less overall than the groups given non-reinforced pretraining.

Nevertheless, the importance of the excitatory summation effect reported here is that it raises an interesting interpretative problem for models of negative patterning. As noted above, the Pearce model does not predict summation after non-reinforced pretraining like that used here. The Rescorla–Wagner model, on the other hand, does predict summation. However, it is not immediately obvious that this summation should have depended upon pretraining.

One effect of pretraining might be to diminish any unconditioned suppressive effects of the stimuli that could prevent excitatory summation from normally occurring. Problematic for this view, however, is the finding in Experiment 2b that group non-differential reinforcement failed to show excitatory summation. As this group had an equal opportunity to habituate to the unconditioned suppressive effects of the stimuli, excitatory summation would have been expected.

Another way of potentially explaining the excitatory summation effect that occurs after discrimination pretraining makes use of the notion of disinhibition (see Aydin & Pearce, 1995). For example, if conditioned inhibition developed to the non-reinforced stimuli during the pretraining phase, then compounding these stimuli might have produced increased responding due to disinhibition. Although this account cannot be ruled out in the present study, this explanation is somewhat unattractive in view of Rescorla's (1997) results demonstrating that excitatory summation in this paradigm does not depend upon inhibition developing to the stimuli.

One tentative conclusion to be drawn here is that discriminative pretraining may be important for an excitatory summation effect to be observed in the negative patterning procedure. This point potentially has interesting theoretical implications. If one were to interpret excitatory summation as reflecting elemental processing of stimulus compounds, then perhaps through a perceptual learning process (see Hall, 1991) discrimination pretraining increases the tendency to treat compound stimuli elementally.

This conclusion is especially interesting because it is opposite to the claim by Redhead and Pearce (1995) that animals solve negative patterning discriminations by treating the compound in a holistic fashion. Support was provided for this claim by the results from experiments that examined the influence of stimulus salience on negative patterning. A series of computer simulations described by Redhead and Pearce (1995) reveals that an elemental model (i.e. the Rescorla–Wagner model with unique cues) predicts that the discrimination between the compound and the more salient element of a negative patterning discrimination will progress more rapidly than that between the compound and the less salient element. The Pearce model (Pearce, 1987, 1994), in contrast, predicts the opposite outcome, principally because there will be less generalization between the compound and the less salient element than between the compound and the more salient element. Experimental results reported by Redhead and Pearce (1995) supported the configural model. Thus, if the data from the present studies are to be interpreted as reflecting elemental processing of the compound, then it is important to examine in the present situation the influence of stimulus salience on the acquisition of the negative patterning discrimination.

It is noteworthy that for the negative patterning discrimination of Experiment 1b subjects in the negative patterning problem experienced greater difficulty discriminating F from the FN compound than they did N. Of course, it is possible that reduced responding to F resulted from the emergence of conditioned rearing or orienting behaviours that competed with magazine responding. It is also possible, however, that the difference is one of associability. If this were the case, then it becomes important to understand which of the two stimuli (N or F) is relatively more salient. If the excitatory summation consistently seen at the outset of negative patterning learning in the present studies reflects elemental processing of the compound, and if these subjects continued to use an elemental strategy to solve the problem, then one would expect to discover that N is more salient than F. The final experiment examined this question using a procedure where competing behaviours were not likely to play a role.

EXPERIMENT 3

The present study investigated the relative salience of the N and F stimuli in an effort to extend the results reported by Redhead and Pearce (1995). Two groups of rats were trained on different instrumental discrimination learning tasks. Rats in group patterning were exposed to an instrumental patterning discrimination in which the correct response in a two-choice discrete-trials procedure was contingent on whether the stimuli were presented individually or in compound. For instance, on F alone and N alone trials, lever pressing would be the correct response, but on FN compound trials, chain pulling would be the correct response. Strictly speaking, this task is neither a negative nor a positive patterning task, because reinforcement could occur on all trials. Nevertheless, it is a patterning discrimination in so far as it requires different responses on compound and element trials. One of the virtues of using an instrumental patterning choice task here is that it effectively eliminates the role of competing behaviours that may be conditioned to either of the two stimuli. If the results in this task were to replicate those observed in the Pavlovian procedures of Experiment 1b, then N should acquire discriminative control more rapidly than should F.

In order to assess the relative salience of N and F, a control group was included. In this group, N and F served as discriminative stimuli in a two-choice discrete-trials instrumental conditional discrimination task in which each stimulus signalled a different correct response alternative. For this group, no compound stimulus trials were included. The degree of stimulus control acquired by each stimulus can be taken as an index of their salience.

Method

Subjects

Sixteen experimentally naive male Sprague-Dawley rats (supplied by Charles River Breeders) weighing approximately 360 g at the beginning of the experiment were used. They were housed and maintained as in Experiment 1.

Apparatus

The apparatus was the same as that used in Experiment 1 with the exception that lever and chain manipulanda were accessible during training sessions. The chain manipulandum was created by suspending a chain through the ceiling from a microswitch mounted on the outer part of the ceiling. The chain was positioned approximately 3 cm to the left of the magazine and about 3 cm away from the front wall on which the magazine was located.

Procedure

The rats were initially given one magazine training session with food pellets. This session was 20 min long during which 20 reinforcers were delivered randomly in time. Following this training, all rats were given lever press training followed by chain pull training with a continuous reinforcement schedule until they had earned approximately 50 reinforcements for each response. Hand shaping was used if necessary. Rats were then given training in different 20-min sessions with each response reinforced according to a variable interval 15-sec schedule. This training continued for 4 days followed by 2 days of concurrent training in which two independent variable interval (VI) 15-sec schedules were in force during each 20-min session.

Instrumental Discrimination Training Following concurrent VI training, the rats were separated into two groups that were matched in terms of their overall rates of responding, as well as their rates of responding to the two responses that were to be designated as “R1” and “R2”. Group patterning received 24 training sessions in which the lever and chain were concurrently available. For every animal, one of the responses (R1+) was reinforced during the singly presented stimuli according to a VI 15-sec schedule while the other response (R2−) was extinguished. During compound trials, this reinforcement contingency was reversed (R1−, R2+). The identity of R1 and R2 was counterbalanced across lever press and chain pull responses for the subjects in each group. Each session contained eight 30-sec trials of each stimulus presented separately and sixteen 30-sec compound stimulus trials. These trials were separated by an ITI of 30 sec. There were four different trial sequences that were used irregularly throughout training.

Group conditional discrimination was trained in the manner just described except that there were no compound trials. One response was reinforced in the presence of N while the other was reinforced in the presence of F. The stimulus-correct response assignments were counterbalanced across subjects in this group. Following Session 12, subjects in this group were trained with a discrimination reversal. Training of the reversal problem continued for an additional 12 sessions.

Results and Discussion

Acquisition of the instrumental patterning discrimination in group patterning is presented in the upper panel of Figure 7. That figure shows R1 and R2 responding in the presence of stimuli F, N, and FN over four-session blocks of training. Due to considerable individual differences in the rates of lever press and chain pull responding, the data have been transformed to elevation ratios. This ratio takes the form A/ (A+B), where A refers to the rate of responding during the stimulus and B refers to the rate of responding in the prestimulus period. Note that in the presence of F and N, R1 was reinforced while R2 was extinguished, but that R2 was reinforced while R1 was extinguished in the presence of FN. Generally, the data show that correct choice responding developed rapidly on FN trials, and more rapidly on N than on F trials. In other words, the discrimination between FN and N emerged sooner than that between FN and F.

These data were statistically analysed with a Stimulus (F, N, FN) × Response (correct vs. incorrect) × Block ANOVA. All main effects and interactions were significant in this analysis. In order to evaluate further the three-way interaction, $F(10, 70) = 2.05$, separate Stimulus × Response ANOVAs were performed for each four-session block following Kirk (1968). These tests revealed significant two-way interactions at Blocks 2, 3, 5, and 6, smallest, $F(2, 15) = 3.96$. Separate repeated measures ANOVAs were then performed on Block 6 comparing correct and incorrect responding in the presence of each stimulus. This analysis revealed significantly more correct than incorrect responding in the presence of N, $F(1, 15) = 12.16$, and FN, $F(1, 15) = 28.08$, but not in the presence of stimulus F.

A separate statistical analysis was performed on the prestimulus response rate data. A Response (R1/ R2) × Block ANOVA revealed a significant Response × Block interaction, $F(5, 35) = 4.52$. This interaction indicates that whereas prestimulus R1 responses increased across training, prestimulus R2 responses decreased. Overall, R1 responses (15.6 responses per min) appeared greater than R2 responses (9.7 responses per min), but this main effect was not significant.

The lower panel of Figure 7 shows data from the acquisition and reversal phases for group conditional discrimination. Subjects were better at selecting the correct response in the presence of stimulus F than in the presence of N during both the acquisition and the reversal phases. The data from the final three blocks of each phase were analysed with a

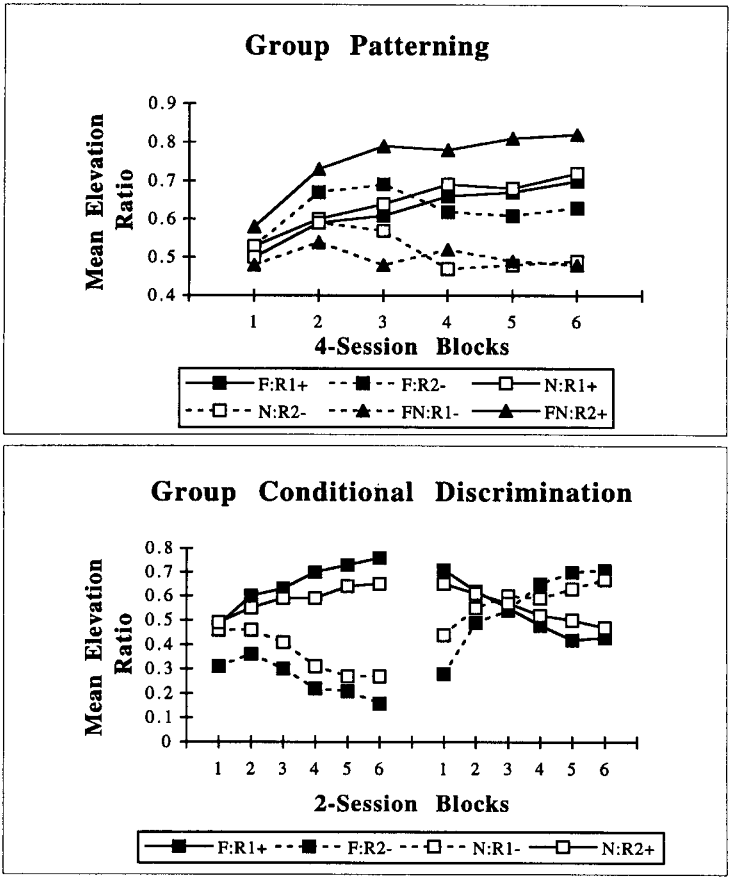


FIG. 7. Mean R1 and R2 responding in the presence of each stimulus across four-session blocks for group patterning and group conditional discrimination in Experiment 3. Elevation ratios were formed by conditionalizing response rates during the stimuli upon prestimulus rates.

Phase (acquisition, reversal) × Stimulus (F, N) × Response (correct, incorrect) × Block (final three blocks of acquisition and reversal) ANOVA. This analysis revealed a significant Stimulus × Response interaction, *F*(1, 7) = 7.66, that did not interact with phase. This result indicates that discriminative responding during F was superior to that during N by the end of each phase.

A separate Response (R1/R2) × Block × Phase ANOVA was performed on the prestimulus response rate data. This analysis revealed a significant main effect of block, *F*(2, 14) = 8.66, as well as a significant Phase × Block interaction, *F*(2, 14) = 3.35. Each of these reflect a general decline in responding over the final three blocks of each phase. Overall, R1 responding (10.8 responses per min) did not differ from R2 responding (12.9 responses per min).

The results from the present experiment were consistent with those from Experiment 1b in demonstrating faster discrimination between FN and N than between FN and F in the patterning task. In addition, the results provided independent evidence that F was more salient than N. This evidence came in the form of F acquiring more stimulus control than N in both the acquisition and reversal phases of the (non-patterning) conditional discrimination task. These results together are consistent with, and thus extend, those reported by Redhead and Pearce (1995) in showing greater difficulty in discriminating the salient than the weak element from the compound in a negative patterning problem.

GENERAL DISCUSSION

The present set of experiments illustrate several interesting characteristics about patterning discrimination learning. First, both negative and to some extent positive patterning learning are facilitated by a pretraining phase in which the elements of the patterning discrimination are reinforced rather than non-reinforced (cf., Forbes & Holland, 1980). Second, pretraining on the discrimination task used here seemed necessary for the occurrence of excitatory summation during the early stage of negative patterning learning. Third, in the negative patterning task discrimination was more difficult between the compound and the more salient element than between the compound and the less salient element.

The present set of studies were aimed at evaluating a fundamental claim of certain connectionist theories that compound stimulus representations change over the course of a patterning discrimination task. This claim was evaluated in several ways. First, unique predictions arising from a specific network model concerning the effects of pretraining upon patterning learning were examined. Second, evidence was sought for the more general claim made by these sorts of models that elemental and configural processes might both be employed at different times in solving patterning tasks. Although the results from the present studies do not provide strong support for the particular network model inspiring these experiments, they do provide some support for the idea that both elemental and configural processes contribute to patterning learning.

In the first experiment the standard feedforward network introduced by Rumelhart et al. (1986) was evaluated. As discussed in greater detail above, that model makes the prediction that reinforced pretraining should facilitate the learning of a positive patterning discrimination, and that non-reinforced pretraining should facilitate the learning of a negative patterning discrimination. The results from Experiments 1a and 1b do not support these predictions. Instead, the results suggest that reinforced pretraining might influence subsequent patterning learning by enhancing the processing and therefore the associability of reinforced stimuli. This effect apparently would have precedence over and above any hypothesized effects that pretraining might have upon the establishment of configural network representations.

The results from Experiments 2 and 3 are of interest as they provide support for the more general claim from connectionist theory that both elemental and configural processes might contribute to patterning learning. As noted above, elemental and configural theories differ in their expectation of excitatory summation during the early portion of negative patterning learning. Excitatory summation is predicted by an elemental but not usually by a configural model. Although Pearce's configural model (Pearce, 1987, 1994) can be made to produce excitatory summation in some settings, as discussed previously it is questionable how applicable it is to the present finding of summation (cf. Darby & Pearce, 1997).

Thus, it is tempting to conclude that the excitatory summation effect seen during the early portion of negative patterning learning reflects elemental processing of the compound stimulus. However, the results of Experiment 3 present serious problems for an elemental model. That experiment demonstrated that the elements used in the negative patterning task were of unequal salience, and that discrimination between the compound and the less salient element was easier than that between the compound and the more salient element. As noted by Redhead and Pearce (1995) this result is precisely that anticipated by the Pearce configural model and is opposite to that predicted by the elemental model of Rescorla and Wagner.

Taking these two results together, it would appear that both elemental and configural processes may contribute to negative patterning learning but that elemental processes are engaged sooner. This conclusion is at odds with models of discrimination learning that start with either elemental or configural assumptions. Moreover, the conclusion suggests that we begin to think seriously about how to incorporate into our models the possibility that both elemental and configural processes contribute throughout acquisition.

The theoretical approach advocated here is a connectionist one. The standard feed-forward model of Rumelhart et al. (1986) did not accurately predict the results from the first set of experiments reported here. Nevertheless, the fundamental assumption of these types of models that internal stimulus representations undergo significant changes throughout acquisition has received some support from the present studies. In particular, the finding that both elemental and configural processes contribute to negative patterning learning is entirely consistent with this framework. One can think, for example, of excitatory summation occurring at the outset of negative patterning because an internal representation of the compound CS has not yet become differentiated from its elements. Once the internal representations of the compound and its elements become differentiated, then a configural representation is said to be learned, and the discrimination problem can be solved.

These possibilities were evaluated in a simulation of the negative patterning task when stimulus salience differs between the elements of the discrimination problem. This simulation, given the problems noted above, used a slightly modified version of the network depicted in Figure 1. The modification consisted of an addition of direct connections between input (CS) and output (US) layers of units to the network depicted in Figure 1. This modification, initially advanced by Schmajuk and DiCarlo (1992; also Schmajuk, Lamoureux, & Holland, 1998), can be thought of as producing an elemental/configural hybrid model. The direct input–output layer connections are an instantiation of the Rescorla–Wagner model, but since differentiated compound representations may emerge at the hidden layer this model also includes a configural component. It is noteworthy that a major difference between this model and the Schmajuk and DiCarlo (1992) model is not in the architecture, but in the learning rules applied to the architecture. Unlike in Schmajuk and DiCarlo (1992), Rumelhart et al.'s (1986) back-propagation learning rule was used throughout the network here.

Figure 8 presents the simulation results of a Pavlovian negative patterning problem when stimulus salience is assumed to differ. Notice that over training both of the results of

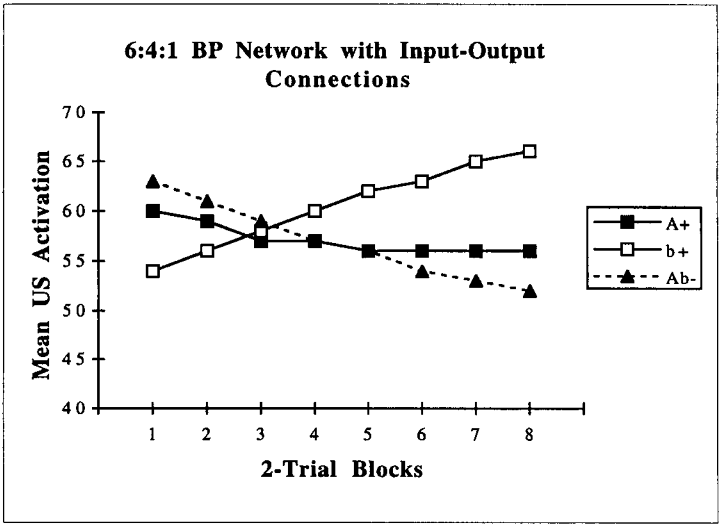


FIG. 8. Simulation results for a negative patterning task in which the stimuli differ in salience. Stimulus A was assumed to be more salient than stimulus b. This task was solved by a 6: 4: 1 network that included direct CS–US connections. Compound summation is seen early, but note that the network learns to discriminate the less salient element from the compound more rapidly than it does the more salient element.

interest reported in the present studies were produced. Summation is seen early, but discrimination between the more salient element and the compound is slow relative to discrimination between the less salient element and the compound. The chief point to be drawn here is not that this successful simulation should persuade the reader of the appropriateness of this particular network model of patterning learning, but rather that the connectionist approach with its assumption of changing internal stimulus representations may lead to models that are more faithful to the data than are existing models that lack this assumption.

In summary, the present set of results seem to point to more ways in which our understanding of patterning learning is limited. Strictly elemental or strictly configural approaches are unlikely to succeed in explaining learning in patterning tasks. Such an understanding may require that we develop models that incorporate assumptions of each of these approaches (see also Kehoe & Graham, 1988). The connectionist network perspective clearly does this, although its success will depend upon how the interaction between elemental and configural processes is conceived.

REFERENCES

Aydin, A., & Pearce, J.M. (1995). Summation in autoshaping with short and long-duration stimuli. *Quarterly Journal of Experimental Psychology, 48B*, 215–234.

Aydin, A., & Pearce, J.M. (1997). Some determinants of response summation. *Animal Learning & Behavior, 25*, 108–121.

Bellingham, W.P., Gillette-Bellingham, & Kehoe, E.J. (1985). Summation and configuration in patterning schedules with the rat and rabbit. *Animal Learning & Behavior, 13*, 152–164.

Darby, R.J., & Pearce, J.M. (1997). The effect of stimulus preexposure on responding during a compound stimulus. *Quarterly Journal of Experimental Psychology, 50B*, 203–216.

Forbes, D.T., & Holland, P.C. (1980). Positive and negative patterning after CS preexposure in flavor aversion conditioning. *Animal Learning & Behavior, 8*, 595–600.

Gluck, M.A., & Myers, C.E. (1993). Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus, 3*, 491–516.

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

Kehoe, E.J. (1988). A layered network model of associative learning: Learning to learn and configuration. *Psychological Review, 95*, 411–433.

Kehoe, E.J., & Graham, P. (1988). Summation and configuration: Stimulus compounding and negative patterning in the rabbit. *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 320–333.

Kirk, R.E. (1968). *Experimental design: Procedures for the behavioral sciences*. Belmont, California: Brooks/ Cole Publishing Company.

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

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207–238.

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., Adam, J., Wilson, P.N., & Darby, R.J. (1992). Effects of discrimination training on responding during a compound conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes, 18*, 379–386.

Redhead, E.S., & Pearce, J.M. (1995). Stimulus salience and negative patterning. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 48B*, 67–83.

Rescorla, R.A. (1972). ”Configural’’ conditioning in discrete-trial bar pressing. *Journal of Comparative and Physiological Psychology, 79*, 307–317.

Rescorla, R.A. (1973). Evidence for the “unique stimulus’’ account of configural conditioning. *Journal of Comparative and Physiological Psychology, 85*, 331–338.

Rescorla, R.A. (1991). Separate reinforcement can enhance the effectiveness of modulators. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 259–269.

Rescorla, R.A. (1997). Summation: Assessment of a configural theory. *Animal Learning & Behavior, 25*, 200–209.

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

Rescorla, R.A., Grau, J.W., & Durlach, P.J. (1985). Analysis of the unique cue in configural discriminations. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 356–366.

Rodger, R.S. (1974). Multiple contrasts, factors, error rate, and power. *British Journal of Mathematical and Statistical Psychology, 27*, 179–198.

Rodger, R.S. (1975). The number of non-zero, post hoc contrasts from ANOVA and error-rate. I. *British Journal of Mathematical and Statistical Psychology, 28*, 71–78.

Rudy, J.W., & Sutherland, R.J. (1992). Configural and elemental associations and the memory coherence problem. *Journal of Cognitive Neuroscience, 4*, 208–216.

Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart, J.L. McClelland, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 1: Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Schmajuk, N., & DiCarlo, J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. *Psychological Review, 99*, 268–305.

Schmajuk, N.A., Lamoureux, J.A., & Holland, P.C. (1998). Occasion setting. A neural network approach. *Psychological Review, 105*, 3–32.

Whitlow, J.W., & Wagner, A.R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science*, *27*, 299–301.

Manuscript received 8 January 1998
Accepted revision received 19 October 1998

APPENDIX A

In models of the sort used here, hidden and output unit activations are given by two equations:

$$\text{Netinput}_{(j)} = \Sigma_i W_{(ij)} * \text{Act}_{(i)} \qquad 1$$

$$\text{Act}_{(j)} = 1/(1 + e^{-\text{netinput}_{(j)}}) \qquad 2$$

Notice that the activation of unit *j* is a logistic function of the weighted sum of activations of those units that are connected to unit *j*. The activation values for input layer units are assumed to take on the value of 1 for active units and 0 for inactive units. Assuming a set of random starting weights (between −0.5 and +0.5) connecting input-to-hidden units and hidden-to-output units, activation of the output unit can be calculated using the equations above. However, weights will need to be adjusted in order to ensure that the network learns to activate appropriately the output unit when different input units are active (i.e. on different trials). In order for weight adjustments to be calculated, error signals must be generated for output and hidden layer units. These error signals will serve as "teaching" values that serve as the activation goal for that unit. The error signal for output units is given by:

$$\delta_{(o)} = (\lambda - \text{Act}_{(o)}) * (1 - \text{Act}_{(o)}) * \text{Act}_{(o)} \qquad 3$$

where λ is the value of reinforcement (1) or non-reinforcement (0). The product of the second and third terms is the derivative of the activation function for the output unit. Inclusion of this derivative is required for weight adjustments to result in reductions in error—that is, for "gradient descent" learning to occur. The error signal for hidden layer units is a weighted portion of this output error:

$$\delta_{(h)} = (1 - \text{Act}_{(h)}) * \text{Act}_{(h)} * \Sigma_o W_{(ho)} * \delta_{(o)} \qquad 4$$

The learning rules are then applied to input-to-hidden unit and hidden-to-output unit connections. The rule governing changes in the latter type of connections is given by:

$$\Delta W_{(ho)} = \text{Act}_{(h)} * \delta_{(o)} * \alpha + \mu * \Delta W_{(ho)} \qquad 5$$

where α is a learning rate parameter. Similarly, the learning rule for changes in input-to-hidden unit connections is given by:

$$\Delta W_{(ih)} = \mathrm{Act}_{(i)} * \delta_{(h)} * \alpha + \mu * \Delta W_{(ih)} \qquad 6$$

In the present simulations the overall speed of learning was facilitated by multiplying a "momentum" parameter (μ) by the change in weight on the previous trial and adding this to the change in weight calculated for the current trial. In the present simulations the values of α and μ were 0.1 and 0.9, respectively.

APPENDIX B

To simulate two discriminations along different sensory continua, similarity was coded in terms of common elements. Thus, pretraining was accomplished with six different input units (A, B, C, D, X, and Y) combined to form four different trial types: AX+, BX−, CY+, DY−. Following near-asymptotic training on this problem, the network was trained on either a positive patterning task in which the stimuli had been reinforced (i.e. AX−, CY−, AXCY+) or one in which the stimuli had been non-reinforced (i.e. BX−, DY−, BXDY+). Similarly, training on the negative patterning task used either the previously reinforced stimuli (i.e. AX+, CY+, AXCY) or the previously non-reinforced stimuli (i.e. BX+, DY+, BXDY−).

The simulated data in Figure 2 average across 10 simulated runs through the problem.

Processus élémentaire et configuré dans l'apprentissage discriminatoire de combinés

Trois expériences utilisant une procédure de conditionnement appétitif chez le rat ont examiné le degré auquel des processus élémentaires et configurés sont impliqués dans les discriminations de combinés positifs et négatifs. La première expérience a démontré que la discrimination des combinés négatifs (et des combinés positifs) fut apprise plus rapidement quand ces discriminations impliquaient des éléments ayant été renforcés auparavant, en comparaison avec des éléments nonrenforcés durant une simple discrimination. L'expérience 2 a révélé un effet de sommation excitatoire au début de l'apprentissage des combinés négatifs attribuable à la phase pré-entraînement. La dernière expérience a démontré un apprentissage discriminatoire plus rapide entre le composé et l'élément le moins manifeste au lieu de celui qui était le plus manifeste dans une tâche de combinés instrumentale. Ces résultats furent interprétés en terme de la théorie connexioniste qui prédit que les représentations interne des stimuli conditionés changent durant le cours d'une discrimination de combinés.

Procesos elementales y configuracionales en el aprendizaje discriminativo de patrón

En tres experimentos se usó condicionamiento apetitivo con ratas para examinar la implicación de los procesos elementales y configuracionales en las discriminaciones de patrón positivo y negativo. El primer experimento demostró que las discrimonaciones de patrón negativo y, hasta cierto punto, positivo se aprendían más rápidamente cuando estas discriminaciones constaban de elementos estimulares que se habían reforzado previamente en comparación con los elementos no reforzados de una discriminación simple. El experimento 2 reveló un efecto de sumación excitatorio durante le primera fase de aprendizaje de patrón negativo que dependía de la discriminación de preentrenamiento. El último experimento demostró un aprendizaje discriminativo más rápido entre el compuesto y el elemento menos saliente que con el más saliente de una tarea de patrón instrumental. Este conjunto de resultados se interpreta en el sentido de que refleja la posibilidad, consistente con la teoría conexionista, de que las representaciones internas de loss estímulos condicionados cambian durante una discriminación de patrón.