Journal of Experimental Psychology:
Animal Behavior Processes
1996, Vol. 22, No. 4, 395-404

0097-7403/96/$3.00

Backward Sensory Preconditioning

Jasper Ward-Robinson and Geoffrey Hall
University of York

In Experiments 1 and 2, rats received initial training in which two neutral events were presented as a serial compound (A→X). Subsequent training with A as a signal for shock was found to endow X with the ability to evoke the conditioned response of suppression. Experiment 2 also showed that responding to X was diminished if, prior to testing, Stimulus A underwent extinction. Two possible mechanisms for these findings are considered: (a) that X elicits responding through the associative chain X–A–shock, and (b) that A activates a representation of X that gains direct associative strength during conditioning with A and loses it during extinction of A. Experiment 3 demonstrated that an X–shock association established after initial A→X training can be extinguished by nonreinforced presentations of A. These results suggest that associatively evoked representations of stimuli can enter into associations.

In standard demonstrations of sensory preconditioning (e.g., Prewitt, 1967), subjects are given an initial phase of training with a serial compound event X–A, neither of the components of which has any marked motivational significance or response-eliciting power. In a second phase of training, one of the elements of the compound (A) undergoes standard Pavlovian training and, by virtue of its association with a motivationally significant unconditioned stimulus (US), comes to evoke an overt conditioned response (CR). In the final, test, phase of the procedure it is demonstrated that Stimulus X is also capable of evoking the CR. This result has been interpreted in terms of the formation of an association between X and A during the first phase of training. The second phase establishes an A–US association so that presentation of the X stimulus at test is able to contact the representation of the US (and thus evoke the CR) by way of the associative chain X–A–US.

In addition, the effect can be found when A and X are presented as a simultaneous compound (e.g., Brogden, 1939; Rescorla & Freberg, 1978); in this case too, an excitatory X–A association could still be formed and might be responsible for the result observed. The remaining temporal arrangement (the *backward* case), in which A precedes the presentation of X, has been little studied; and to the extent that this procedure is less likely to generate the excitatory X–A association, it might be supposed that it would be unlikely to yield a sensory preconditioning effect. Indeed, with one exception, studies that have used backward pairings (i.e., A–X) in the first phase of training have failed to find any effect (Brown & King, 1969; Coppock, 1958; Tait, Marquis, Williams, Weinstein, & Suboski, 1969).

The exception is an experiment by Silver and Meyer (1954), who gave rats extensive initial training in which a 1-s presentation of Stimulus A was followed after 0.5 s by a 1-s presentation of X. An avoidance response subsequently established to A showed positive transfer to X (comparison being made with control groups given exposure to A alone, X alone, or no pretraining.) This result is theoretically intriguing, and if substantiated it might require us to rethink our standard interpretation of the phenomenon of sensory preconditioning. Unfortunately, however, it is not clear that the effect obtained by Silver and Meyer (1954) depended on the *backward* pairings of A and X in Phase 1. The intertrial interval used in this phase of training was such that the interval between the offset of X on one trial and the onset of A on the next was less than 4 s. In these conditions a forward excitatory association between X and A could well have been formed. A satisfactory demonstration of backward sensory preconditioning requires a procedure in which the consequences of possible associations between events presented on different trials are controlled for. In the first experiment described below we attempted to provide this. In subsequent experiments we began an analysis of the mechanisms that might be responsible for the effect.

Experiment 1

This experiment used rats as the subjects, the conditioned emotional response procedure in the second (conditioning) phase, and the within-subject design summarized in Table 1. All rats received Phase 1 training in which presentations of the target stimulus (X) were preceded immediately by presentations of Stimulus A. In Phase 2, A was established as a signal for shock and in the test phase any tendency of X to evoke the CR of suppression was assessed. Comparison was made with a control stimulus, Y. This stimulus received Phase 1 training equivalent to that given to X, being presented the same number of times and preceded reliably by another event, B. The procedure differed only in that B was nonreinforced in Phase 2. AX and BY trials occurred intermixed during Phase 1 so that X preceded the presentation of

This work was supported by a grant from the Biotechnology and Biological Sciences Research Council. We thank C. M. T. Bonardi for helpful discussion on an earlier version of this article.

Correspondence concerning this article should be addressed to Jasper Ward-Robinson or Geoffrey Hall, Department of Psychology, University of York, York YO1 5DD, United Kingdom. Electronic mail may be sent via Internet to jwr1@york.ac.uk or gh1@york.ac.uk.

Table 1
Experimental Designs

Phase 1	Phase 2	Phase 3	Test
	Experiment 1		
A→X	A→shock	NA	X
B→Y	B→no shock	NA	Y
	Experiment 2		
Group Ext			
A→X	A→shock	A→no shock	X
B→Y	B→no shock	B→no shock	Y
Group VI			
A→X	A→shock	—	X
B→Y	B→no shock	—	Y
	Experiment 3		
A→X	X→shock	A→no shock	X
B→Y	Y→shock		Y

Note. A, B, X, and Y represent visual and auditory stimuli. In Experiments 1 and 3, all subjects experienced all of the trial types shown. In Experiment 2, there were two groups, differing in their Phase 3 training. Group Ext (extinction) received nonreinforced stimulus presentations in this phase; for Group VI (variable interval), no stimulus presentations were scheduled.

the A stimulus (on the next trial) as often as it preceded the presentation of B. Similarly, Stimulus Y was followed, after the intertrial interval, as often by A as by B. It was possible, then, that X might form forward associations with both A and B, but the same would be true of control Stimulus Y. A difference between X and Y in the ability to evoke suppression on the test could not, therefore, be attributed to across-trial forward associations formed in Phase 1.

Method

Subjects

The subjects were 32 female hooded (Lister) rats with a mean ad-lib weight of 160 g (range = 145–175 g). These rats had been used in a teaching demonstration of instrumental learning but were naive with respect to the stimuli and apparatus used in the current experiment. They were maintained at 80% of their free-feeding weights by being fed a measured amount of food after the last training session of each day. The rats were housed in pairs in a colony room illuminated from 8:00 a.m. to 8:00 p.m.

Apparatus

Four identical Skinner boxes (Campden Instruments Ltd., Loughborough, UK) were used. Each was housed in a sound-attenuating and light-proof shell. The boxes were equipped with a recessed food tray to which 45-mg pellets could be delivered. The tray was covered by a sprung, transparent plastic flap (6 cm high by 5 cm wide), which was hinged at the top. Pushing against this flap actuated a microswitch, closure of which was recorded as a response. The standard response levers were retracted through the course of the experiment. The floor was made from stainless steel rods that could be electrified by a Campden Instruments Ltd. shock generator (Model 521C) and shock scrambler (Model 521S). There were two loudspeakers, one mounted on the wall opposite the food tray, the other on the wall opposite the box's door. These were used to present a 2-Hz train of clicks at 80 dB and a white noise at 90 dB (Scale A). The exhaust fan which served to ventilate the chamber generated a background noise level of 65 dB. Illumination was provided by a 1.5-cm-diameter, 3-W jewel light (rated for 24 V but operated at 15V) mounted 14.5 cm above the base of the magazine tray. Offset of this light constituted the *dark* stimulus. The ceiling of the box consisted of a sheet of translucent plastic, and above this was positioned a 30-W strip-light. This was used to provide the stimulus referred to as *light,* below. Events were controlled and recorded with a BBC microcomputer (Model B) that used a version of BASIC.

Procedure

Pretraining. Initially the rats were given two 40-min sessions of magazine training in which pellets were delivered according to a variable-time 60-s schedule. After the rats had learned to push aside the magazine flap and retrieve food pellets, pushing the flap was trained as an instrumental response. Subjects were required to earn 25 pellets according to a continuous reinforcement schedule in the third pretraining session and to respond on a variable interval (VI) 60-s schedule in the next session. This and all subsequent sessions were 40 min in duration. Responding was maintained on the VI 60-s baseline throughout the rest of the experiment. The experiment was run between 10:00 a.m. and 5:00 p.m. daily. Each of the eight squads of rats was run at a consistent time of day during all phases of the experiment.

Phase 1. Over the next six sessions, the rats received A–X and B–Y trials. Two trials of each type were presented during each session in quasi-random sequence; the sequence was constrained to ensure that, over all six sessions, a given trial type was followed equally often by a trial of the same type as by a trial of the other type. For half of the subjects, A was the clicker and B the dark stimulus; for the remainder, these stimulus assignments were reversed. For half of the subjects in each of these groups, X was the light and Y was the white noise; for the remaining rats, the arrangement was the reverse. In other experiments the light and 90-dB noise have proved to be highly salient to rats. We therefore anticipated that the use of light and noise as test stimuli might be of benefit in producing a backward sensory preconditioning effect. The stimulus durations, chosen on the basis of pilot work, were 30 s for A and B and 1 s for X and Y. Onset of the second stimulus on a trial immediately followed the offset of the first. The intertrial interval (ITI), measured from the offset of one stimulus to the onset of the next, was 462 s.

Phase 2. In each of the next two sessions, all subjects received two trials, one a presentation of A followed by a 1.0-mA, 0.5-s footshock, the other a nonreinforced presentation of B. Half of the subjects received the sequence A, B, B, A (over the two sessions); the remainder received the sequence B, A, A, B. Responding was recorded during both the conditioned stimulus (CS) and the 30-s pre-CS periods. The ITI was 790 s. There followed two sessions of baseline recovery training in which responding was reinforced on the VI 60-s schedule and no other events were programmed to occur.

Testing. In each of the three test sessions there were three presentations of X and three of Y. The order of trials was random with the constraint that no more than two trials of the same sort could occur in succession and that for half of the subjects the first test trial was with Stimulus X, whereas for the others it was with Stimulus Y. In order to allow a reasonable sample of behavior to be obtained, we increased the duration of each stimulus presenta-

tion to 30 s. Responding was also recorded during the 30-s pre-CS periods. The ITI was 322 s.

Results and Discussion

All subjects readily learned to respond on the VI 60-s schedule and maintained responding throughout the A–X and B–Y trials of Phase 1. The reinforced training of Phase 2 resulted in a suppression of responding in the presence of Stimulus A; responding was maintained in the presence of the nonreinforced Stimulus B. A suppression ratio of the form $a/(a + b)$, where a is the response rate during the stimulus and b the rate during the prestimulus period, was computed for each trial during this phase. The mean suppression ratio governed by Stimulus A was .45 on Trial 1, but it fell to .26 on the second trial. Stimulus B produced mean suppression ratios of .44 and .41 on these trials. An analysis of variance (ANOVA) with trial and stimulus type as the factors was performed on these data. It revealed a main effect of trial, $F(1, 31) = 15.35$, no main effect of stimulus type, $F(1, 31) = 1.73$, but a significant interaction between these factors, $F(1, 31) = 7.32$ (in all statistical tests reported in this article, a rejection criterion of $P \leq .05$ was adopted). The source of this interaction was examined with a test of simple main effects. The difference in suppression governed by A and B on the second trial only was found to be reliable, $F(1, 31) = 6.15$.

The results of principal interest are presented in Figure 1, which shows the suppression evoked by each test stimulus (X and Y) on each of the three test sessions. A suppression ratio was calculated for each trial, and the scores for each trial type were averaged, thus producing a suppression ratio for each of the stimuli on each session. As the figure shows, both stimuli initially evoked a moderate amount of suppression. With continued nonreinforced testing, suppression was lost, so that by the final session both stimuli evoked ratios of around .50. Before this point was reached, however, it is evident that Stimulus X (that paired in Phase 1 with the stimulus that was reinforced during Phase 2) elicited more suppression than did Stimulus Y. An ANOVA with session and stimulus as the factors was conducted on the data summarized in the figure. This yielded no main effect of stimulus, $F(1, 31) = 2.17$, but a main effect of session, $F(2, 62) = 37.53$, and a significant interaction between these factors, $F(2, 62) = 4.19$. We performed an examination of simple main effects to locate the source of this difference. A simple main effect of stimulus was found to be reliable on Session 2, $F(1, 31) = 4.86$. Neither difference on Days 1 or 3 was reliable (smallest $p = .13$). Group mean pre-CS response rates from these test sessions are presented in Table 2. It can be seen that mean pre-CS rates from X and Y trials were similar and did not change in any systematic way across the three test sessions. An ANOVA with stimulus and session as factors was performed on these data and produced no reliable statistics (smallest $p = .07$).

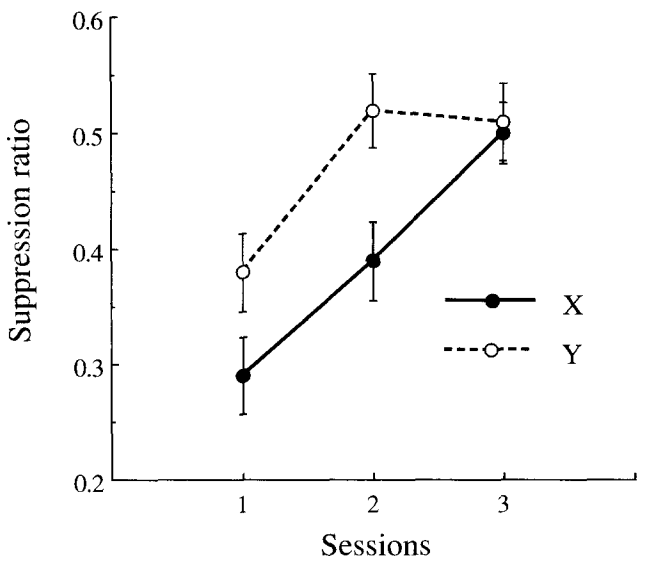

Figure 1. Experiment 1: Group mean suppression ratios for Stimuli X and Y in the test phase. In Phase 1, X had been presented signaled by A and Y by B. In Phase 2, A was paired with shock. Error bars indicate standard error of the mean.

That both Stimuli X and Y evoked suppression at the start of the test phase is open to more than one explanation. One possibility is that both elicit some unconditioned suppression, the relatively brief exposure given to these stimuli during Phase 1 being insufficient to allow habituation of this effect. Alternatively, the associative strength acquired by Stimulus A during Phase 2 could have generalized to both X and Y. But the *difference* in performance to X and Y can be attributed only to the fact that X had been signaled in Phase 1 by the subsequently reinforced Stimulus A, whereas Y had been signaled by the nonreinforced B. Evidently in these circumstances the suppression governed by A is more likely to generalize to X than to Y.

This finding constitutes an instance of sensory preconditioning in that pairing two stimuli was found to allow a CR subsequently acquired by one of the pair to be controlled by the other. But it constitutes a demonstration of a backward version of the phenomenon in that during the first phase of training the test stimulus (X) was presented after rather than before the stimulus (A) that was to be trained as a CS in Phase 2. It is true that on some occasions during Phase 1, X occurred prior to a presentation of A. This allowed the possibility that a forward, X–A association might be formed; but the formation of such an association cannot explain the result obtained (the difference on test between Stimuli X and Y) because both X and Y had the opportunity to form such across-trial associations. Rather, this result must be a consequence of the fact that X had been preceded by A and Y by B in the training trials of Phase 1.

Experiment 2

The effect demonstrated in Experiment 1 is, from certain points of view, both surprising and theoretically interesting. It seemed important, therefore, to attempt to confirm its reliability, and our first aim in Experiment 2 was to replicate the basic result of Experiment 1.

Table 2
Test Session Pre-CS Response Rates (Responses per Minute)

	Stimulus X		Stimulus Y	
Test session	*M*	*SD*	*M*	*SD*
Experiment 1				
Session 1	18.08	9.13	14.73	7.98
Session 2	15.52	8.45	15.04	9.02
Session 3	16.92	9.55	16.71	11.45
Experiment 2				
Group VI				
Trial 1	23.00	14.35	19.13	17.66
Trial 2	19.38	21.36	16.38	17.71
Trial 3	22.13	18.61	20.00	13.27
Group Ext				
Trial 1	18.50	12.03	21.75	16.36
Trial 2	17.25	17.45	17.88	12.01
Trial 3	21.38	13.40	17.00	15.16
Experiment 3	19.45	8.46	17.97	9.69

Note. Stimuli X and Y followed the pre-CS period. CS = conditioned stimulus; VI = variable interval; Ext = extinction.

We also sought to begin analysis of the factors that might contribute to the backward sensory preconditioning effect by investigating the extent to which the effect depends on the associative status of the stimulus reinforced in Phase 2. In standard procedures for the study of sensory preconditioning (i.e., experiments that use forward pairings or simultaneous presentations of the stimuli in Phase 1 of training), it has been shown that the effect depends on the status of the stimulus trained as a CS in Phase 2 of the experiment. Nonreinforced presentations of the CS, interpolated between the conditioning and test phases of the sensory preconditioning procedure, reduce or eliminate the ability of the test stimulus to evoke the CR (e.g., Rescorla & Cunningham, 1978; Rizley & Rescorla, 1972). Our second goal in the present experiment was to determine whether backward sensory preconditioning might similarly be abolished when the reinforced stimulus undergoes extinction before testing.

One group of subjects (Group Ext) received training similar to that given to the subjects in Experiment 1 except for the insertion of a phase of extinction of the CSs between Phase 2 conditioning and testing. Subjects in Group VI were treated identically except that for them this additional phase consisted simply of baseline training. This latter group might be expected to show backward sensory preconditioning, confirming the reliability of the effect and also serving as a control against which the effects of the extinction treatment given to Group Ext could be assessed. The design of the experiment is summarized in Table 1.

Method

The subjects were 32 female hooded Lister rats with a mean ad-lib weight of 160 g (range = 130–190 g). They were from the same stock and maintained in the same way as the rats used in Experiment 1. The apparatus was that used in Experiment 1. In other experiments from our laboratory we have found light and dark to be especially well matched in the unconditioned suppression they elicit; accordingly, these events were used as Stimuli X and Y in the present experiment. The clicker and white noise served as A and B. In Experiment 2 the intensity of the white noise was reduced to 80 dB, matching that of the clicker. All other stimulus intensities remained the same as in Experiment 1.

Pretraining of the instrumental baseline was conducted in the same way as was described for Experiment 1. Next, in Phase 1, all subjects received A–X and B–Y trials just as in the previous experiment. Stimulus assignments were counterbalanced as in Experiment 1. Phase 2 consisted of shock-reinforced A trials and nonreinforced B trials. Because Experiment 1 revealed no effect of the trial sequence used, all rats in this experiment received the same sequence of trials in this stage: A, B, B, A.

For Phase 3, the rats were assigned to one of two equal-sized groups: Group Ext and Group VI. The counterbalanced subgroups were equally represented in each of these groups. Over the following 8 days rats in Group Ext received nonreinforced presentations of A and B. In each of the first six 40-min sessions, four A and four B trials were presented in a random sequence, with the constraint that no more than two trials of a given type could occur in succession. Trials were separated by a 240-s ITI. On Days 7 and 8 the session duration was increased to 60 min, and 10 presentations of each stimulus were given, the ITI being reduced to 143 s. Thus these subjects received a total of 44 nonreinforced presentations of each of the stimuli previously trained as CSs. As a consequence of a programming error, Group VI received the same treatment as Group Ext on the first day of this phase, but thereafter these subjects experienced only baseline VI training with no stimuli being presented. This means that subjects in Group VI received four nonreinforced trials with each CS. This was, however, insufficient to produce any marked decline in the ability of these stimuli to evoke suppression.

Finally, all subjects received a single test session containing three presentations of X and three of Y. Half of the rats in each counterbalanced subgroup received the trial sequence X, Y, Y, X, Y, X; the remainder received Y, X, X, Y, X, Y. In details not specified here, the procedure was the same as in Experiment 1.

The experiment was run in two replications, the second being started immediately after the first had been completed. The replications were identical except for the way in which the stimuli were counterbalanced: For Group Ext, Stimulus X was dark in Replication 1 and light in Replication 2; for Group VI, this arrangement was reversed. Both replications were run between 1:00 and 5:00 p.m. daily.

Results and Discussion

As in Experiment 1, the instrumental baseline was readily established and maintained throughout training. The introduction of A–X and B–Y trials in Phase 1 had little effect on responding, and no data are presented for this phase. The shock-reinforced trials of Phase 2 established suppression of responding in the presence of Stimulus A. One can assess the effects of this phase of training most conveniently by examining performance on the first day of Phase 3, a day on which all subjects received four presentations of A and four of B. Suppression ratios, computed after pooling scores across all trials of a given type, showed that, for both groups, A evoked suppression and B did not. For Group Ext, the mean ratios were .11 to A and .51 to B; for Group VI the equivalent scores were .17 and .55. An ANOVA revealed a

main effect of stimulus, $F(1, 30) = 93.28$, but no significant effect of group, $F(2, 30) = 2.80$, and no significant interaction, $F < 1$. The groups did not differ in their baseline levels of responding; pre-CS rates during Phase 2 were 12.9 and 12.0 responses per minute for Groups Ext and VI, respectively. These rates did not differ significantly, $F < 1$.

Repeated exposure to A in Phase 3 was effective in producing extinction of suppression in Group Ext. On the final day of this phase the mean suppression ratio to A was .54, and that to B was .50. These scores did not differ significantly, $F < 1$.

The data of central interest, group mean suppression ratios for each test trial, are displayed in Figure 2. The results for Group VI are shown in the top panel of the figure. Both stimuli evoked some suppression that diminished over the course of testing, but Stimulus X (the stimulus that was signaled by Stimulus A in Phase 1) elicited more suppression than did Stimulus Y. For Group Ext, on the other hand, there was no clear difference between X and Y, suppression to both these stimuli being approximately equivalent to that seen to Stimulus Y in Group VI. An ANOVA with group, stimulus, and trial as the factors was carried out on the data summarized in Figure 2. This yielded a significant effect of trial, $F(2, 60) = 36.15$, and of stimulus, $F(2, 60) = 10.27$, and an interaction between group and stimulus that fell just short of the conventional level of significance, $F(1, 30) = 3.94$, $p < .06$. No other effects were reliable (smallest $p = 0.11$). A separate ANOVA conducted on the data for Group VI demonstrated that there was a significant main effect of stimulus, $F(1, 15) = 14.00$, and a significant effect of trial, $F(2, 30) = 11.45$. These factors did not interact, $F < 1$. An equivalent analysis for Group Ext revealed only a main effect of trial, $F(2, 30) = 31.44$; for stimulus type, $F < 1$, and for the interaction, $F(2, 30) = 1.06$. Mean pre-CS response rates from the test session are presented in Table 2. It can be seen that the rates were similar for both groups across all trials. Statistical examination of these data supported this; an ANOVA with group, stimulus, and trial as factors revealed no reliable effects (smallest $p = .23$).

The results for Group VI exactly match those of Experiment 1 and thus demonstrate the reliability of the backward sensory preconditioning effect observed in the previous experiment. The new finding of Experiment 2 is that the difference between the test stimuli is abolished when the CR previously governed by Stimulus A is extinguished prior to testing. Both stimuli still evoked some suppression at the beginning of the test, but only at a level similar to that shown to Stimulus Y in Group VI. This level of suppression, we suggest, reflects the subjects' unconditioned response to the test stimuli; the extra degree of suppression shown to Stimulus X in Group VI is a conditioned effect that depends on the establishment and maintenance of the CR to X's associate, Stimulus A. Thus, like forward and simultaneous cases of sensory preconditioning, the effect produced by the backward training procedure is abolished by extinction of the test stimulus's associate prior to testing.

The explanation for this extinction effect offered in the case of standard sensory preconditioning procedures assumes that the ability of the test stimulus to evoke a CR

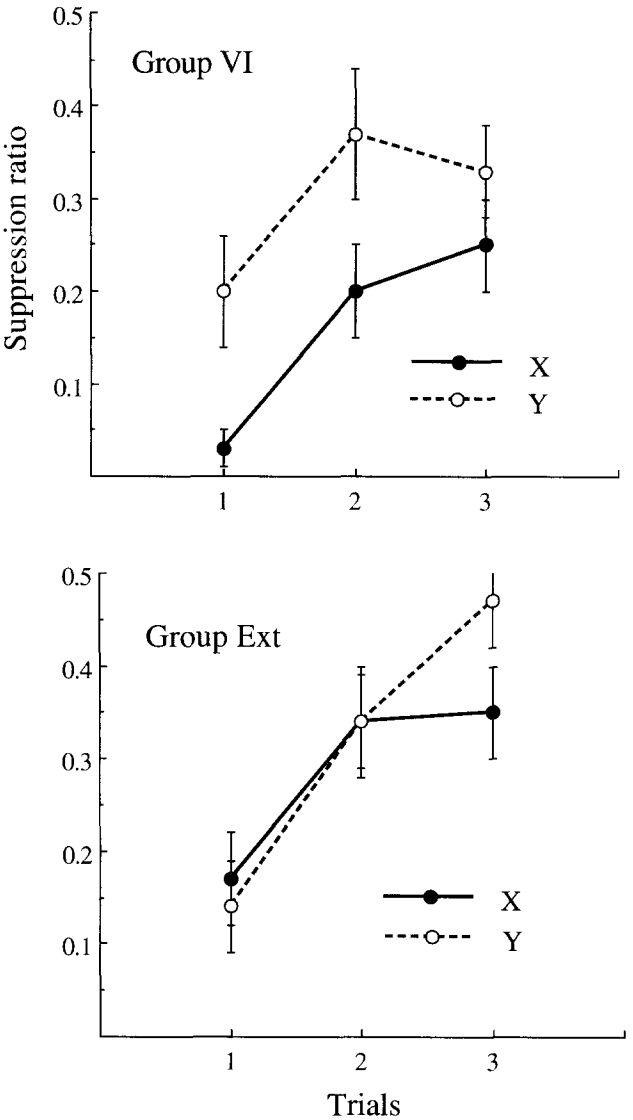

Figure 2. Experiment 2: Group mean suppression ratios for Stimuli X and Y in the test phase. For all subjects, training was the same as described for Experiment 1 except that for Group Ext (extinction) a phase of extinction of Stimulus A preceded the test. Group VI (variable interval) received no stimulus presentations during this phase, but continued to respond on a variable-interval baseline schedule. Error bars indicate standard error of the mean.

depends on the operation of an associative chain. Conditioning with the CS (Stimulus A) establishes, it is assumed, an association between representations of A and the US which allows A to evoke a CR. Phase 1 training is presumed to establish an excitatory association between the stimulus to be tested (X) and Stimulus A. Thus, on test, X is able to contact the US representation and elicit a CR by way of the chain X–A–US. The loss of the sensory preconditioning effect when extinction of Stimulus A occurs prior to the test is taken to reflect a break in the last link of this chain. And the sensitivity of the backward sensory preconditioning

effect to extinction of A seems, at first sight, to imply that this phenomenon too requires an intact X–A–US chain. There are, however, problems with this analysis.

The associative-chain interpretation of sensory preconditioning requires that, in the first phase of training, Stimulus X become established as an excitatory CS, signaling Stimulus A. For standard forward and simultaneous training procedures, this assumption creates no problems. For our Phase 1 training procedure, however (in which presentations of A preceded those of X), it requires us to accept that excitatory backward conditioning occurred. Studies of orthodox classical conditioning show that backward pairings can indeed result in excitatory learning in some circumstances (see, e.g., Spetch, Wilkie, & Pinel, 1981) and thus the associative-chain account of our backward sensory preconditioning effect is a possibility. It should be acknowledged, however, that backward pairings have been shown to produce excitation only in a rather restricted set of conditions (e.g., when rather few training trials are given; Heth, 1976), and these conditions were not especially well met in our Phase 1 procedure. We consider next an alternative explanation for the effect seen in our experiments that does not require us to accept that excitatory backward conditioning occurs with these procedures.

Experiment 3

It is not contentious to assume that our Phase 1 procedure would allow the formation of a (forward) excitatory association between A as the CS and Stimulus X as the US. The existence of such an association allows the possibility that backward sensory preconditioning might occur, not by way of an associative chain, but because of the formation of a direct link between the test stimulus and the US representation. One effect of the A–X association formed in Phase 1 will be that the representation of X will be activated during Phase 2 when A is presented and paired with the reinforcer. Since the representation of X is paired with the US it is at least possible that an X–US association might form. The X–US association will thus generate a CR when X is actually presented in the test phase. (See Rescorla & Cunningham, 1978; Rescorla & Freberg, 1978, for the application of this notion to the case of simultaneous sensory preconditioning.)

The results of Experiment 2 may seem to present a problem for this analysis: If the backward sensory preconditioning effect depends on an X–US association formed in Phase 2, why should changes in the strength of the A–US link (such as will be brought about by extinction trials in Group Ext) reduce the size of this effect? In fact, a solution to this apparent problem follows directly from the basic proposal that animals can learn about associatively activated event representations. If such learning can occur and is responsible for the formation of an X–US association in Phase 2, then it might also be expected to go on during the extinction phase of Experiment 2. In particular, nonreinforced presentations of A will activate X and thus could result in the representation of X being paired with the absence of shock; that is, this training could result in extinction of the X–US association and thus a loss of responding on the test.

Holland and Forbes (1982) provided direct evidence for such a representation-mediated extinction effect in a study of conditioned flavor aversion. In their experiments rats were first trained with a tone as a signal for the availability of sucrose; then an aversion was established to the sucrose by following its consumption with the injection of a toxin. A phase of exposure to the tone on its own was found to result in a reduction in the strength of the sucrose aversion. Holland and Forbes concluded that activation of the representation of sucrose (by presenting the event that had been established as a CS for sucrose in the first phase of training) in the absence of the toxin could produce extinction of the conditioned aversion.

Although this study established the reality of mediated extinction, it did so for procedures very different from those used in the experiments reported here. We thought it important, therefore, to attempt to demonstrate the phenomenon using essentially the same procedures as in Experiments 1 and 2. This would provide further evidence in favor of the general position that associatively activated event representations can take part in associative learning, and it would provide support for the more specific suggestion that the result seen in Group Ext of Experiment 2 was a consequence of a process of mediated extinction.

All of the rats in this experiment received Phase 1 training identical to that given in Experiment 2 (see Table 1), training that can be assumed to establish an A–X association. In the second stage of the present experiment, however, X (rather than A) was paired with the shock US. The question of interest was whether subsequent presentations of A, in the absence of the US, would produce mediated extinction of the CR governed by X. Phase 3 therefore consisted of A-alone presentations and was followed by a test phase in which suppression to X was assessed. Comparison was made with Stimulus Y, which had received equivalent Phase 1 training (having been signaled by B) and had been paired with shock in Phase 2. It differed only in that no B-alone presentations were given in Phase 3, thus providing no opportunity for mediated extinction. Our expected result, therefore, was that Y would elicit more suppression than X.

Method

The subjects were 16 female hooded Lister rats with a mean ad-lib weight of 170 g (range = 155–200 g). They were maintained at 85% of their free-feeding weights. The apparatus and stimuli were identical to those used in Experiment 2.

After pretraining on the instrumental baseline, all subjects received six sessions of Phase 1 training consisting of presentations of A–X and B–Y. The detailed procedures were exactly as described for Experiment 2. On each of the next four sessions (Phase 2), all subjects received a single shock-reinforced trial occurring halfway through the 40-min session. The CS duration was 30 s. Half of the subjects received the sequence X, Y, Y, X, and half received the sequence Y, X, X, Y. Phase 3 consisted of three sessions, each containing six presentations of Stimulus A alone separated by an ITI of 322 s. Each of the four test sessions that

followed contained three presentations of X and three of Y. The stimulus durations remained at 30 s. The experiment was run daily between 1:00 and 5:00 p.m. In any details not specified here, the procedure was the same as in the previous experiments.

Results and Discussion

One rat became ill and died part way through the experiment, and the results reported here are thus for the remaining 15 subjects.

No data are presented for Phase 1. Phase 2 training successfully established suppression to the reinforced stimuli. The group mean suppression ratio for the final trial with Stimulus X in this stage was .05; the equivalent score for Stimulus Y was .04. There was little suppression to Stimulus A in Phase 3. Daily suppression ratios were calculated using each rat's pooled pre-CS and CS rates on that day. Group means over the 3 days of Phase 3 were .50, .53, and .57. It may be noted (see Table 1) that this phase of training is formally equivalent to the test phase of a standard experiment on sensory preconditioning; some suppression to A might therefore have been expected (at least early in the phase, before the occurrence of extinction brought about by repeated testing). Accordingly, the very first trial of this phase was looked at separately. The group mean suppression ratio on this trial was .38, a result consistent with the possibility that A came to this phase with a modest ability to evoke the CR established to X in Phase 2. In the absence of an appropriate control condition, however, it is impossible to assert with any confidence that the suppression exhibited on this trial constitutes an instance of sensory preconditioning.

The results of primary interest come from the final phase in which X and Y were tested. For each subject, response rates were collapsed over all 12 test trials with a given stimulus, and a suppression ratio was calculated. The group mean scores, displayed in Figure 3, show that X controlled less suppression than Y. These scores differed significantly, $F(1, 14) = 5.47$. This is what would be expected if the treatment given in Phase 3 of training allowed mediated extinction to occur—if activation of the representation of X in the absence of shock brought about extinction of the X→shock association that had been established in Phase 2. The pre-CS response rates that were used to calculate these suppression ratios are summarized in Table 2. It can be seen that the rates differed only slightly during X and Y trials. An ANOVA with stimulus only as a factor found this difference to be unreliable ($p = .17$).

The results of this experiment demonstrate that a mediated extinction effect can be found with the present training procedures; they thus add weight to the suggestion that mediated extinction might be responsible for the result found for Group Ext in Experiment 2. In that experiment too, nonreinforced presentations of A might be expected to allow extinction of any strength possessed by its associate, X. According to this analysis, the only important difference between the two experiments is that in Experiment 3, X acquired its strength by direct reinforcement, whereas in Experiment 2, associative mediation was involved in the initial acquisition as well as the extinction. The proposition that associatively activated representations can be the subject of associative learning thus provides a coherent account for the entire pattern of results.

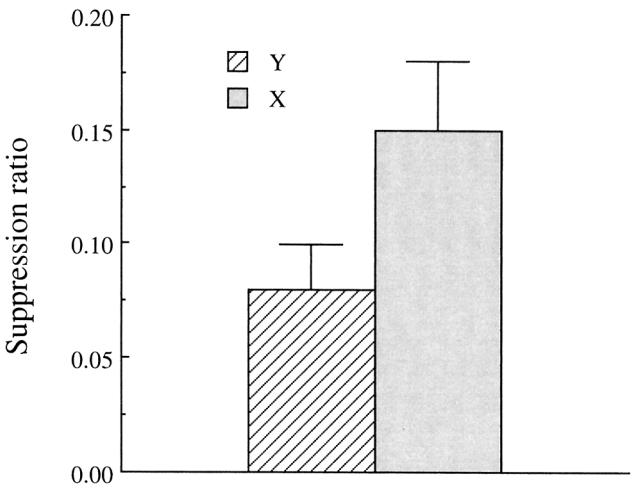

Figure 3. Experiment 3: Group mean suppression ratios for Stimuli X and Y, pooled over all test sessions. Both X and Y had been established as signals for shock in Phase 2; in Phase 3, which immediately preceded the test, all subjects received nonreinforced presentations of an event that had been used to signal the occurrence of X in the first phase of training. Error bars indicate standard error of the mean.

General Discussion

Experiments 1 and 2 constitute, to the best of our knowledge, the first unambiguous demonstrations of backward sensory preconditioning. They show that a CR established to one stimulus (A) will transfer to another stimulus (X) when rats have experienced a preliminary phase of training in which A has signaled the occurrence of X. The only previous demonstration of such an effect, provided by Silver and Meyer (1954), is arguably inadequate. As we have noted, these authors used a short ITI during the first phase of training, which raises the possibility that forward conditioning (between the X stimulus that terminated one trial and the A stimulus that started the next) might have been responsible for their effect. In our experiment we made use of a much longer ITI and a within-subject design that eliminated any contribution from such across-trial forward associations as might still be formed. We conclude, therefore, that the ability of X to evoke the CR depends on the A→X training of Phase 1.

Given that our training parameters appear to generate a reasonably robust backward sensory preconditioning effect, we need to ask why previous studies (Brown & King, 1969; Coppock, 1958; Tait et al., 1969) should have failed to find one. Any answer must be speculative at this stage. It is worth noting, however, that all of the earlier studies gave rather more Phase 2 trials (a minimum of 10) than were given in our procedure. Rescorla (1983) noted that the

standard sensory preconditioning effect can be abolished by extended Phase 2 training, and the same may be true of the backward version of the phenomenon.

We have considered two possible explanations for backward sensory preconditioning. One is that backward excitatory conditioning occurs during the A–X trials of Phase 1. X would then be able to activate a representation of A which, after Phase 2 training, would allow X to make contact with a representation of the US in the test phase. The second is that forward conditioning occurs in Phase 1 and that the associatively activated representation of X forms a direct association with the US in Phase 2. Both of these possibilities can accommodate the finding of Experiment 2 that presentations of A alone, prior to the test, will attenuate the backward sensory preconditioning effect. According to the first, the A–US link will be weakened, and with it the effectiveness of the X–A–US chain on which the effect is held to depend. According to the second, presentation of A activates the X representation in the absence of the US, and thus mediated extinction of the X–US association will occur.

Although both accounts can explain the results of Experiments 1 and 2, only the latter can explain those of Experiment 3. This experiment established the reality of the mediated extinction effect with the use of procedures modeled on those of Experiment 2 and thus gives plausibility to the suggestion that this process was at work in the previous experiment. The mediated learning account may thus be preferred on grounds of parsimony because it is able to deal with the results of all three experiments. We cannot assert, however, that backward excitatory (X–A) learning made no contribution to the backward sensory preconditioning effect observed in Experiments 1 and 2.

Note the parallel between these experiments and a set of studies by Holland (e.g., 1990) on the phenomenon he referred to as representation-mediated conditioning. In one of these experiments (Holland, 1981) rats were given Phase 1 training in which a tone signaled the presentation of a distinctively flavored sucrose pellet. In a second phase, presentation of the tone was followed by a nausea-inducing injection of lithium chloride. These subjects were found to shun the pellets (compared with control rats that had received unpaired presentations of the events in one or another of the two phases). This finding is formally identical to the backward sensory preconditioning effect of the present experiments in that the CR established to the tone in Phase 2 transferred to an event that had been signaled by that tone in Phase 1. Holland's (1981) interpretation of his finding was that the initial appetitive training phase endowed the tone with the ability to activate a representation of the sucrose pellet that became associated with the illness induced by lithium chloride in the second phase.

Holland (1981) also provided direct evidence against the alternative interpretation—the possibility that the sucrose is shunned on test because backward excitatory conditioning occurs in Phase 1 and that this, along with Phase 2 conditioning, establishes the associative chain of sucrose–tone–illness. He included a control condition consisting of subjects given training designed explicitly to establish the sucrose pellet as a signal for the tone in Phase 1 (i.e., for these subjects a tone sounded just after the pellet had been delivered). If the rejection of sucrose on test in the experimental group was the product of a sucrose–tone–illness associative chain, then the rejection of sucrose by these control rats that received direct sucrose–tone pairings should be at least as strong as in the experimental group. In fact no such aversion was observed in these rats, which supports Holland's (1981) interpretation of the effect as being the result of representation-mediated learning.

When taken together, Holland's findings and our own present a reasonably strong case for the idea that rats can learn about associatively activated representations of stimuli. This may have implications for our theorizing about associative learning. One of the most successful formal models of this process (Wagner, 1981) supposes that representations of stimuli may hold two states of activity which differ in their properties. After a stimulus has been presented, its representation will be in its fullest state of activation, which Wagner (1981) termed A1. When in the A1 state, a representation may enter into associations with other active representations. However, according to Wagner, a representation in the less active state of A2 will be unable to act as a CS in associative learning. One way in which a stimulus representation comes to enter into A2 activity is by being activated by an associate. In our experiments, an associatively activated stimulus (X) appears to be able to become a CS (for shock). To this extent, our results (along with those of Holland, 1981, 1990) suggest the need to reconsider some aspects of Wagner's (1981) theory (see Holland, 1983), in particular to allow that a stimulus representation in the A2 state can acquire associative strength.

Such a revision of Wagner's (1981) theory could provide a satisfying account both of the results reported here and of Holland's own work (e.g., Holland, 1981). However, an alternative account provided by Matzel, Held, and Miller (1988) requires consideration. Matzel et al. (1988) proposed a "temporal coding hypothesis" (see also Barnet, Arnold, & Miller, 1991) according to which the pairing of a CS and a US results not only in the formation of an association between representations of these stimuli but also in the encoding of temporal information about the relationship between them during the conditioning episode. Matzel et al. (1988) claimed that it is the nature of this temporal information rather than the association per se that governs conditioned responding. For example, it is suggested that the typically poor conditioned responding observed following simultaneous Pavlovian conditioning (e.g., Smith, Coleman, & Gormezano, 1969) is not due to a deficit of associative learning but instead occurs because, in this procedure, the CS does not bear a predictive relationship to the US. A rat given such training will show little anticipatory responding because it has knowledge that the CS occurs with, rather than predicts, the US.

The hypothesis assumes that temporal information may be integrated across different phases of training, and this makes it possible to anticipate backward sensory preconditioning under circumstances in which the test stimulus provides predictive information about the onset of the US. A

recent experiment (Experiment 2) reported by Cole, Barnet, and Miller (1995) examined this suggestion. In that experiment, rats first received presentations of a 5-s tone followed by a 5-s click stimulus. They were next given tone–shock pairings, with a 5-s trace interval between offset of the tone and occurrence of the shock. Finally, fear responding to the click was assessed. The temporal coding hypothesis predicts that integrating temporal and associative information would lead rats to show conditioned responding to the click on test. And indeed, rats given this training showed more fear responding than a control group given similar treatment but with contiguous presentation of tone and shock in the second phase of training.

Note that the rats in the control condition in the experiment by Cole et al. (1995) received training that was formally identical to that used in the present Experiments 1 and 2 to demonstrate the backward sensory preconditioning effect. We may assume, therefore, that the effect was generated in both groups in the experiment by Cole et al. (1995) and take their results as showing that the magnitude of the effect is enhanced when the shock-reinforced stage of training involves a trace interval. Such an outcome can be accommodated by the proposal that backward sensory preconditioning depends on the acquisition of associative strength by the associatively activated representation of the test stimulus. We do not know the time course over which presentation of the tone generates A2 activity in the representation of the click, and it is thus entirely open to us to assume that concurrent activation of the click and shock representation is better achieved when there is a delay between tone and shock than when tone and shock are contiguous.

The issue for the temporal coding hypothesis is that of explaining why there should be any backward sensory preconditioning effect at all when, as in our experiments, the shock is presented immediately after Stimulus A in Phase 2 of training. According to this hypothesis (see Cole et al., 1995, pp. 148–149), integration of the two phases of training would supply the information that the onset of X and the onset of the shock occur at the same point in time (on the offset of A). Stimulus X would not therefore bear the predictive relationship to shock that the temporal coding hypothesis supposes to be necessary for conditioned responding to occur. In this form, at least, the hypothesis is disconfirmed by the present results. It should be acknowledged, however, that the hypothesis needs only minor amendment in order to accommodate our findings. Although in our experiments shock immediately followed the offset of A, it is possible that full activation of the US representation takes some time to develop after the onset of the shock. If so, then X might nonetheless be construed as standing in a predictive relationship to the critical reinforcing event.

Clearly it is necessary to conduct more experimental work on the temporal relations between stimuli that are necessary for producing the backward sensory preconditioning effect. An unusual feature of our procedures (which were otherwise quite standard) was the use of a brief event as the X stimulus during the A–X pairings of Phase 1. This arrangement was chosen simply because our first indication of the backward sensory preconditioning effect came as the chance result of another study, directed at a different issue and involving the use of brief stimuli. We simply carried over to the present experiments the stimulus durations used in the earlier experiment. We do not know whether these durations are of importance, but it is interesting to note that in Holland's (1981) experiment, which, we have argued, constitutes a further case of backward sensory preconditioning, the event used as X (the delivery of a sucrose pellet) might be thought to be of relatively short duration. In future work we intend to investigate, for our procedure, the effects of changing the duration of the stimuli.

References

Barnet, R. C., Arnold, H. M., & Miller, R. R. (1991). Simultaneous conditioning demonstrated in second order conditioning: Evidence for similar associative structure in forward and simultaneous conditioning. *Learning and Motivation, 22,* 253–268.

Brogden, W. J. (1939). Sensory pre-conditioning. *Journal of Experimental Psychology, 25,* 323–332.

Brown, C. P., & King, M. G. (1969). Backward sensory preconditioning: The ineffectiveness of the procedure under optimal conditions. *Australian Journal of Psychology, 21,* 55–58.

Cole, R. P., Barnet, R. C., & Miller, R. R. (1995). Temporal encoding in trace conditioning. *Animal Learning & Behavior, 23,* 144–153.

Coppock, W. J. (1958). Pre-extinction in sensory preconditioning. *Journal of Experimental Psychology, 55,* 213–219.

Heth, C. D. (1976). Simultaneous and backward fear conditioning as a function of number of CS–US pairings. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 117–129.

Holland, P. C. (1981). Acquisition of representation-mediated conditioned food aversion. *Learning and Motivation, 12,* 1–12.

Holland, P. C. (1983). Representation-mediated overshadowing and potentiation of conditioned aversions. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 1–13.

Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition, 37,* 105–131.

Holland, P. C., & Forbes, D. T. (1982). Representation mediated extinction of conditioned flavor aversions. *Learning and Motivation, 13,* 454–471.

Matzel, L. D., Held, F. D., & Miller, R. R. (1988). Information and expression of simultaneous and backward associations: Implications for contiguity theory. *Learning and Motivation, 19,* 317–344.

Prewitt, E. P. (1967). Number of preconditioning trials in sensory preconditioning using CER training. *Journal of Comparative and Physiological Psychology, 64,* 360–362.

Rescorla, R. A. (1983). Effect of separate presentation of the elements on within-compound learning in autoshaping. *Animal Learning & Behavior, 11,* 439–446.

Rescorla, R. A., & Cunningham, C. L. (1978). Within-compound flavor associations. *Journal of Experimental Psychology: Animal Behavior Processes, 4,* 267–275.

Rescorla, R. A., & Freberg, L. (1978). The extinction of within-compound flavor associations. *Learning and Motivation, 9,* 411–427.

Rizley, R. C., & Rescorla, R. A. (1972). Associations in second-order conditioning and sensory preconditioning. *Journal of Comparative and Physiological Psychology, 81,* 1–11.

Silver, C. A., & Meyer, D. A. (1954). Temporal factors in sensory

preconditioning. *Journal of Comparative and Physiological Psychology, 47,* 57–59.

Smith, M. C., Coleman, S. R., & Gormezano, I. (1969). Classical conditioning of the rabbit's nictitating membrane at backward, simultaneous and forward CS–US intervals. *Journal of Comparative and Physiological Psychology, 69,* 226–231.

Spetch, M. L., Wilkie, D. M., & Pinel, J. P. J. (1981). Backward conditioning: Reevaluation of the empirical evidence. *Psychological Bulletin, 89,* 163–175.

Tait, R. W., Marquis, H. A., Williams, R., Weinstein, L., & Suboski, M. D. (1969). Extinction of sensory preconditioning. *Journal of Comparative and Physiological Psychology, 69,* 170–172.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Erlbaum.

Received November 13, 1995
Revision received January 4, 1996
Accepted May 15, 1996 ■