

Retrospective revaluation effects in the conditioned suppression procedure

Ian Shevill and Geoffrey Hall

University of York, York, UK

In four experiments using the conditioned suppression procedure, rats received initial reinforced training with two compound stimuli, AX and BY, each compound consisting of one auditory and one visual element. After a second phase of training consisting of nonreinforced presentations of A, the suppression governed by X and Y was tested. In Experiment 1 X evoked slightly less suppression than Y (a mediated extinction effect). This outcome was obtained when the auditory cues served as X and Y (Experiment 1a), when the visual cues served as X and Y (Experiment 1b), and when the number of nonreinforced presentations of A was increased (Experiment 1c) from 18 to 216. In Experiment 2, however, in which the initial training was given with serial compounds (i.e., $A \rightarrow X$ and $B \rightarrow Y$) X evoked more suppression than Y (a recovery-from-overshadowing effect). It was argued that extinction of A engages two learning processes, one increasing the effective associative strength of its associate (X) and one reducing it, and that the balance between these two depends on the specific conditions of training.

The experiments reported here all make use of a version of the following experimental design. The subjects (rats) were trained initially with a compound stimulus (AX) followed by reinforcement. One of the stimulus elements (A) was then presented alone in the absence of the reinforcer. The effect of this treatment on the conditioned response (CR) governed by element X was then tested. All of the three possible outcomes of this treatment (an increase, an attenuation, and no change in the magnitude of the CR) can be predicted by current theories of conditioning, and, on occasion, all three have been obtained. The aim of the present experiments was to attempt to establish the factors that determine which effect will occur, and thus to allow choice among the theories.

According to standard, elemental theories of associative learning, such as the Rescorla–Wagner (1972) model, extinction of A should, in itself, be without effect on the CR governed by X. The presence of X during the initial phase of training with the AX compound can be expected to result in overshadowing so that the associative strength acquired by X would be less than that produced by conditioning of X alone. But once it has been trained, such

Correspondence should be addressed to G. Hall, Department of Psychology, University of York, York, YO10 5DD, UK. Email: GH1@york.ac.uk

This work was supported by a studentship awarded to Ian Shevill by the UK Biotechnology and Biological Sciences Research Council. We thank Jasper Ward-Robinson for much helpful discussion.

associative strength as X may have acquired will be independent of subsequent changes in the strength of A. In an extensive series of experiments using an appetitive conditioning procedure, Holland (1999) has found just this result. Configural theories of conditioning (e.g., Pearce, 1987, 1994) appear to make the same prediction as elemental theories. The configural approach treats the compound, AX, as a unique configure, and any responding controlled by X alone is attributed to generalization from AX to X. It might be thought, then, that the effects of the extinction treatment given to A would generalize to AX and in this way reduce the amount of generalized responding controlled by X. It should be noted, however, that formal theories of configural learning (e.g., Pearce, 1987; see also Wagner & Brandon, 2001) are constrained to argue that generalized associative strength does not itself generalize further. In particular, for this case, inhibitory learning occurring to A may generalize to AX, but it will not generalize further to X. The response to X (although a product of generalization from AX) will not be influenced by treatment given to A.

That nonreinforced presentations of A should increase the magnitude of the CR governed by X is predicted by two quite different theories. Dickinson and Burke (1996) have adopted the notion (proposed by Holland, 1981, 1990; see also Hall, 1996; Konorski, 1967) that learning can occur when the representation of the conditioned stimulus (CS) is activated associatively rather than by direct presentation of the stimulus itself. In the present experimental procedure, initial training with the AX compound can be expected to establish a within-compound association, which will allow the representation of X to be activated associatively on the A-alone trials. Since A was paired with the unconditioned stimulus (US) in initial training, the representation of the US will also be activated associatively on these trials. Dickinson and Burke have proposed that the cooccurrence of two associatively activated representations will lead to the formation (or strengthening) of an excitatory association between them, and thus, in this case, that the strength of the CR to X should be increased (see also Van Hamme & Wasserman, 1994). The same prediction can be derived from the *comparator* theory of Miller and his colleagues (e.g., Blaisdell, Bristol, Gunther, & Miller, 1998; Miller & Matzel, 1988). According to this account, training with the AX compound will not only establish associations between each element and the US, but it will also establish A as a comparator for X. The magnitude of the CR evoked by X when presented alone is held to be determined, in part, by the associative strength of its comparator. Extinction of A, which will reduce the strength of the comparator, will allow the strength governed by X to evoke a more vigorous CR. Such a result, sometimes referred to as a recovery-from-overshadowing effect, has been obtained in studies of human contingency judgements (e.g., Dickinson & Burke, 1996; Larkin, Aitken, & Dickinson, 1998) and in experiments using rats and the conditioned suppression procedure (e.g., Kaufman & Bolles, 1981; Matzel, Shuster, & Miller, 1987; Miller, Barnet, & Grahame, 1992). Results recently reported by Blaser, Couvillon, and Bitterman (this issue 2004, Experiment 3) are consistent with the possibility that this effect may also be observed in an invertebrate species (the honeybee).

That the extinction treatment given to A might attenuate the CR governed by X is again predicted by two quite different accounts. When Holland (e.g., 1990) proposed that an associatively activated CS representation might undergo associative change, he suggested that such a representation would function in much the same way as a directly activated representation. In this case, therefore, X, being associatively activated by A in the absence of the US should undergo extinction and should thus lose strength as a consequence of the A-alone trials

(the reverse of the result predicted by Dickinson & Burke's, 1996, analysis). Results consistent with this analysis have come from studies of flavour-aversion learning (Holland & Forbes, 1982), appetitive classical conditioning (Holland & Ross, 1981), and the conditioned suppression procedure (Ward-Robinson & Hall, 1996). This outcome (sometimes referred to as mediated extinction) can also be predicted, at least for some versions of the experimental design under consideration, by a slightly modified version of standard associative theory. According to this, although the CR governed by X on test will be largely determined by the strength of the X-US association, it is possible that other associations might play a role. In particular, training with a simultaneous AX compound will have established an X-A association, and, since A will also be associated with the US, the associative chain X-A-US could also make a contribution to the ability of X to contact the US representation and thus evoke the CR. Extinction trials with A will reduce the effectiveness of the A-US association, reduce the contribution made by the associative chain, and thus reduce the magnitude of the CR to X. Rescorla and Cunningham (1978) have obtained just this result in an experiment using the flavour-aversion procedure.

In order to resolve the important theoretical issues raised by this disparate set of experimental results it is necessary to identify the factors that are critical in determining which outcome will be obtained. This is difficult to do with the current set of data—an experiment demonstrating retrospective revaluation in human contingency judgement differs in so many ways from one demonstrating mediated extinction in flavour-aversion learning in rats that it is difficult to know where to start. It may be noted, however, that in one experimental procedure (conditioned suppression), both mediated extinction and recovery from overshadowing have been observed (by Ward-Robinson & Hall, 1996, and by Miller and colleagues, e.g., Miller et al., 1992, respectively). Accordingly we thought it sensible to concentrate on this procedure in an attempt to isolate the relevant variables. Given that the mediated extinction effect has already been demonstrated in this laboratory (in the study by Ward-Robinson & Hall) our intended strategy was to first generate an example of recovery from overshadowing using our own conditioned suppression procedure but with detailed parameters based on those used by Miller et al. (1992). We then intended to investigate the various ways in which these parameters differed from those used by Ward-Robinson and Hall, in the hope of identifying the factor or factors that might be effective in converting recovery from overshadowing into the mediated extinction effect. As will become evident, the experimental results actually obtained required us to deviate from our intended programme.

EXPERIMENT 1A

One of the clearest demonstrations of the recovery-from-overshadowing effect is provided by Miller et al.'s (1992) Experiment 3. They used a within-subject design in which the rats were trained initially with two compound stimuli, AX and BY, each associated with shock reinforcement. After a phase in which A was presented alone, the animals were tested with X and Y, and it was found that X elicited greater suppression (of a water-licking response) than did Y. This within-subject design has the advantage of allowing an assessment of the effects of extinguishing A in animals that all receive the same experience of the stimuli and the same amount of exposure to the experimental context. In the present experiment we attempted to replicate this result in our own version of the conditioned suppression procedure. Some

details were different—like Ward-Robinson and Hall (1996) we used a food-reinforced baseline response—but in other respects, what we took to be the essential features of the original experiment were reproduced here.

Method

Subjects

The subjects were 16 male hooded (Lister) rats whose mean ad lib weight was 526 g (range: 480–555 g). The animals had previously been used in an experiment testing flavour preferences but they were naive to the stimuli and procedures used here. Throughout the experiment the rats were housed in pairs in home cages made of opaque white plastic, 35 × 22 × 19 cm. These had a roof of wire mesh that held food and a water bottle; a layer of wood shavings covered the floor. The home cages were kept in a large colony room with a 12-hr light/12-hr dark schedule. Experimental sessions occurred daily during the light phase of the schedule.

Apparatus

Four identical Skinner boxes, supplied by Paul Fray Ltd. (Cambridge, UK), were used. Each was housed in a sound- and light-attenuating shell equipped with an exhaust fan, serving to ventilate the chamber and generating a background noise level of 65 dB. The boxes were equipped with a recessed food tray to which 45-mg pellets could be delivered. A sprung, transparent plastic flap (4 cm high by 4.5 cm wide) covered the tray and 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 Cantab (Cantab Pharmaceuticals, Cambridge, UK) shock generator and scrambler. A loudspeaker mounted on the roof opposite the food tray was used to present a 2-kHz tone with a mean intensity of 78 dB (measured at the food tray). The tone was pulsed: 0.5 s on/ 0.5 s off. A relay was used to present a click train at a rate of 10 per second at an intensity of 80 dB (measured at the food tray). Illumination was provided by a 1.5-cm diameter jewel light with an 8-W bulb, positioned centrally on the wall 14.5 cm above the base of the food tray, which was dimmed by passing its current through a 120- Ω resistor. Turning off this light constituted the stimulus to be referred to as dark. Illumination of an overhead jewel light positioned in the centre of the roof constituted the light stimulus referred to below. This light was 1.5 cm in diameter and was equipped with an 8-W bulb, which was dimmed by passing its current through a 120- Ω resistor. Events were controlled and recorded with a BBC microcomputer (Model B) that used a version of BASIC.

Procedure

In the pretraining phase, the rats received three 40-min sessions of magazine training in which pellets were delivered according to a variable-time 60-s schedule and during which they learned to push aside the magazine flap and retrieve food pellets. Pushing the flap was then trained as an instrumental response. Subjects were required to earn 25 pellets according to a continuous reinforcement schedule in the fourth pretraining session and to respond on a variable-interval (VI) 30-s schedule in the next session. The next four sessions employed a VI 60-s schedule. These and all subsequent sessions were 40 min in duration. Responding was maintained on the VI 60-s baseline throughout the rest of the experiment.

In each of the four sessions of Phase 1 of training, the rats received four trials of one type (i.e., presentations of either the AX or the BY compound) each of which was followed by a 1.0-mA, 0.5-s footshock. Following the procedure used by Miller et al. (1992, Exp. 3), simultaneous compounds were used, and

the duration of each trial was 10 s. Half of the subjects received AX → US trials on Session 1 and BY → US trials on Session 2. The remainder received the opposite arrangement. Sessions 3 and 4 were a repeat of Sessions 1 and 2. For half of the subjects A was the light, and B was the dark stimulus; for the remainder these stimulus arrangements were reversed. For half of the subjects in each of these groups, X was the clicker and Y the tone; for the remaining rats, the arrangement was the reverse. The intertrial interval (ITI), measured from the offset of one stimulus to the onset of the next, was 472 s. Responding was recorded during stimulus presentations and also during the 10-s stimulus-free period (the preCS period) that preceded each trial.

In each of the next three sessions (Phase 2), all subjects received six nonreinforced presentations of A. Trials were separated by an ITI of 334 s. The stimulus duration remained at 10 s.

The test phase consisted of four test sessions in which the suppression of instrumental responding governed by X and Y was assessed. Three trials of each stimulus were given on each of these sessions. 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 behaviour to be obtained, the duration of each stimulus presentation was increased to 30 s. Responding was also recorded during the 30-s preCS periods. The ITI was 317 s.

Results and discussion

Conditioning was assessed in terms of a suppression ratio of the form $a/(a+b)$, where a is the response rate during the CS, and b is the response rate during the corresponding preCS period.

All subjects initially acquired responding on the VI 60-s schedule but with the introduction of shock-reinforced trials in Phase 1, responding declined in many and was completely suppressed in some rats. This made it impossible to monitor trial-by-trial the acquisition of conditioned suppression during this phase. Accordingly the preCS and CS scores for each subject were pooled across all trials of a given type in order to compute a single suppression ratio for each compound. The group means were .09 for AX and .09 for BY.

Baseline responding recovered during Phase 2. For each animal, responding was pooled over all six trials on a given day, and a suppression ratio was calculated. The substantial suppression governed by Stimulus A was lost over the course of nonreinforced presentations; the group mean ratio score was .23 on Day 1, .34 on Day 2, and .39 on Day 3. An analysis of variance (ANOVA) conducted on these data confirmed the reliability of the effect of session, $F(2, 30) = 4.92$. (Here and elsewhere a significance level of $p < .05$ was adopted.)

The results of principal interest are those from the test sessions in which X and Y were presented. Suppression ratios were calculated from the total preCS and CS responding for each subject over all 12 trials with a given stimulus. The mean response rates during the preCS periods were 8.17 responses per min for trials with X, and 7.78 responses per min for trials with Y. These rates did not differ significantly ($F < 1$). The group mean ratio scores shown in Figure 1 (upper panel) indicate that suppression was somewhat more substantial in the presence of Y than of X; that is, Phase 2 presentation of A, far from enhancing the degree of suppression governed by X, appeared to decrease it. Rather than producing recovery from overshadowing, the experimental procedure appears to have produced the result previously referred to as mediated extinction—a loss of conditioned responding in the target stimulus as a result of extinction of its associate.

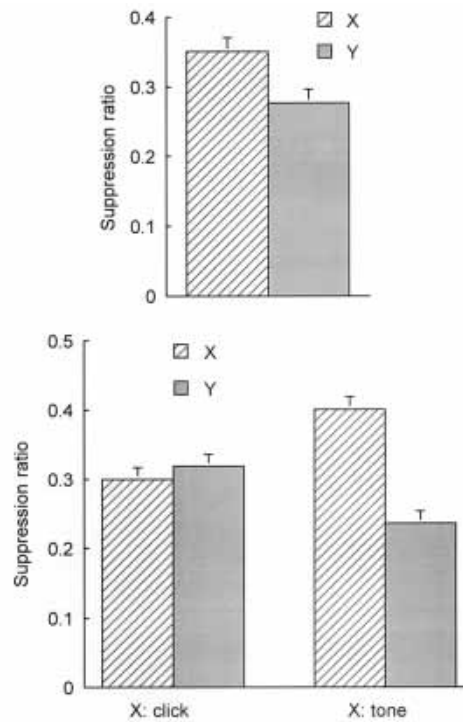


Figure 1. Experiment 1a. Group mean suppression scores for Stimuli X and Y in the test phase (upper panel). The lower panel shows the means for the counterbalanced subgroups, which received either the click as X and the tone as Y, or vice versa. All animals had previously received reinforced trials with the compounds AX and BY followed by nonreinforced presentations of A. Error bars represent within-subjects standard error, computed on scores adjusted for variation between subjects (Bakeman & McArthur, 1996).

Statistical analysis lends support to this conclusion. An ANOVA conducted on the scores summarized in the upper panel of Figure 1 showed that the difference between the means fell just short of the conventional level of significance, $F(1, 15) = 3.10, p < .10$. It was noted, however, that the means presented in the upper panel of Figure 2 (p. 338) obscured a substantial difference between the counterbalanced subgroups in the experiment (i.e., between those animals that received the click as X and the tone as Y, or vice versa). In particular the click proved to be more effective at evoking suppression than was the tone. This is evident in the results presented in the lower panel of the figure, which presents the mean scores for X and Y broken down according to subgroup.

In both subgroups the suppression governed by the click stimulus was more than that governed by the tone, but the difference was enhanced when the intrinsically less suppressive stimulus (the tone) served as Stimulus X. An ANOVA was conducted on the data presented in the lower panel of Figure 1, with subgroup and stimulus (X or Y) as the variables. There was no significant effect of subgroup ($F < 1$), and the main effect of stimulus fell short of significance $F(1, 14) = 4.30, p < .10$, but there was a significant interaction between these variables, $F(1, 14) = 6.79$. An analysis of simple main effects showed that the difference between X and Y

was not significant ($F < 1$) in the subgroup given the click as Stimulus X but was reliable, $F(1, 14) = 10.95$, in the subgroup given the tone as X. This pattern of results is best interpreted as reflecting the influence of two factors—a difference between tone and clicker in their ability to evoke suppression and a loss of suppression by a test stimulus when its associate had been extinguished in Phase 2.

EXPERIMENT 1B

In spite of the fact that the procedures and parameters used in Experiment 1a were modelled on those used by Miller et al. (1992, Exp. 3) in their demonstration of recovery from overshadowing, quite the opposite effect was produced. But although the parameters used here were very similar to those used by Miller et al. they were not identical. Our next step, therefore, was to examine these procedural differences to determine which might be responsible for the discrepant results.

A feature of the parameters used by Miller et al. (1992) in their demonstration of recovery from overshadowing was that the target stimuli used on test (i.e., X and Y in the present notation) were intrinsically less effective at evoking conditioned suppression than were their associates (A and B). In contrast, previous work conducted in this laboratory indicated that the stimuli we used as X and Y were rather more salient than those used as A and B. In the present experiment, therefore, the visual cues (previously used as Stimuli A and B) were used as X and Y, and the auditory cues were used as A and B. In all other respects the design and procedure remained the same as those described for Experiment 1a.

Method

Subjects and apparatus

The subjects were 16 male hooded (Lister) rats whose mean ad lib weight was 495 g (range: 445–560 g). The subjects had been used in an unrelated experiment prior to this one but they were naive to the stimuli and procedures used. The apparatus was that used in Experiment 1a. With the exceptions specifically noted below, the procedure followed was identical to that used in the previous experiment.

Procedure

The rats were first trained to perform the baseline instrumental response. They then received Phase 1 training consisting of reinforced trials with the AX and BY compounds. A and B were the clicker and the tone; X and Y were the light and dark stimuli. The counterbalancing of the stimuli and the order of their presentation was the same as those in Experiment 1a. Phase 2 consisted of nonreinforced presentations of the auditory A stimulus.

Since the stimuli used as X and Y were less effective at evoking suppression than those used in Experiment 1a, we reduced the number of test trials. There were two test sessions each containing two trials with each stimulus. The order of trials was X, Y, Y, X for half the subjects and Y, X, X, Y for the remainder. As with Experiment 1a the duration of each stimulus presentation was increased to 30 s. The ITI was 456 s.

Results and discussion

To an even greater extent than was evident in Experiment 1a the introduction of shocks in Phase 1 resulted in a marked suppression of responding both during the CS and during preCS periods. So few animals showed preCS responding on sessions after the first in this phase that we were unable to compute meaningful suppression ratios. Baseline responding recovered during Phase 2; by the final session of this phase, the mean suppression ratio was .24.

For the test phase, suppression ratios were calculated after pooling the preCS and CS scores for each subject over all four trials with a given stimulus. The preCS rate for trials on which X was presented was 15.63 responses per min; the rate for trials with Stimulus Y was 17.09 responses per min. These rates did not differ reliably, $F(1, 15) = 2.61$. The group mean ratio scores shown in Figure 2 (upper panel) indicate very little difference between X and Y in the suppression they controlled. There was certainly no sign of recovery from overshadowing—indeed, if anything, the result is again in the direction of mediated extinction in that suppression to X was slightly less than that to Y. An ANOVA conducted on the overall means

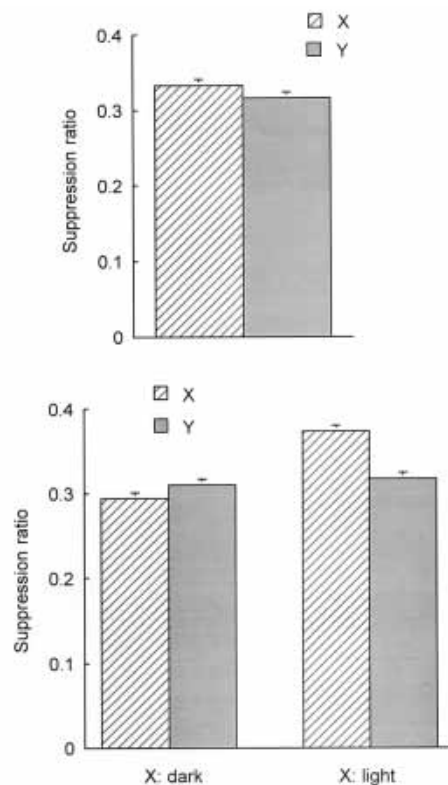


Figure 2. Experiment 1b. Group mean suppression scores for Stimuli X and Y in the test phase (upper panel). The lower panel shows the means for the counterbalanced subgroups, which received either the dark as X and the light as Y, or vice versa. All animals had previously received reinforced trials with the compounds AX and BY followed by nonreinforced presentations of A. Error bars represent within-subjects standard error, computed on scores adjusted for variation between subjects.

presented in the upper panel of Figure 2 showed the difference between them to be nonsignificant, $F(1, 15) = 1.06$. But, as in Experiment 1a, the overall means obscured a substantial difference between the counterbalanced subgroups in the experiment (i.e., between those animals that received the dark stimulus as X and the light as Y, or vice versa). In particular the dark proved to be more effective at evoking suppression than was the light. This is evident in the results shown in the lower panel of Figure 2, which presents the mean scores for X and Y, broken down according to subgroup. In both subgroups the suppression governed by the dark stimulus was greater than that governed by the light, but the difference was enhanced when the intrinsically less suppressive stimulus (the light) served as Stimulus X. An ANOVA was conducted on the data presented in the figure with subgroup and stimulus (X or Y) as the variables. There was no significant effect of subgroup, $F(1, 14) = 1.59$, and no significant effect of stimulus type, $F(1, 14) = 1.46$, but there was a significant interaction between these variables, $F(1, 14) = 6.70$. An analysis of simple main effects showed that the difference between X and Y was not significant ($F < 1$) in the subgroup given the dark as Stimulus X but was reliable, $F(1, 14) = 7.21$, in the subgroup given the light as X. As before, this pattern of results suggests the operation of two factors—a difference between light and dark in their ability to evoke suppression and a loss of suppression by a test stimulus when its associate had been extinguished in Phase 2.

The results of this experiment are thus much the same as those of Experiment 1a. Although the evidence of mediated extinction is less clear here than in the previous experiment, there is certainly no sign of the recovery-from-overshadowing effect. We may conclude that reversing the relative salience of the two sets of cues is not enough to produce the effect and that some other factor must be involved. Experiment 1c explores one possibility.

EXPERIMENT 1C

A feature of the procedure used by Miller and his colleagues in their studies of recovery from overshadowing has been the substantial amount of extinction given in Phase 2 to the associate of the target stimulus. In the demonstration of the effect provided by Miller et al. (1992, Exp. 3) there were 216 Phase 2 extinction trials (as opposed to the 18 trials used in Experiments 1a and 1b). Miller has suggested (R. R. Miller, personal communication) that the number of Phase 2 trials may be a critical variable, and that the recovery-from-overshadowing effect will only be observed when extensive Phase 2 training is given. The present experiment was conducted to test this suggestion. The procedure was in all major respects identical to that described for Experiment 1b except that the number of nonreinforced presentations of Stimulus A in Phase 2 was increased from 18 to 216.

Method

Subjects and apparatus

The subjects were 16 experimentally naive male hooded (Lister) rats with a mean ad lib weight of 443 g (range: 400–515 g) at the start of the experiment. The rats were naive to the stimuli and procedures used. They were maintained in the same fashion as rats of the previous experiments.

Procedure

After the instrumental baseline response had been established, all animals received Phase 1 training consisting of reinforced AX and BY trials. The stimuli were the same as those used in Experiment 1b except that the intensity of the light stimulus was increased (the resistor was reduced to 100 Ω) in an attempt to match the salience of the two visual cues (X and Y) that would be presented on test. In addition the shock intensity was reduced to 0.5 mA in the hope of avoiding the disruption of baseline responding that was produced by the stronger shock intensity used in the previous experiments. In all other respects Phase 1 training was identical to that given in Experiment 1b.

The procedure for Phase 2 was closely modelled on that used by Miller et al. (1992, Exp. 3). In each of the six sessions, all subjects received 36 nonreinforced presentations of A. Trials were separated by an ITI of 136 s. Stimulus duration was kept to 10 s. Sessions were increased in length to 90 min. Technical limitations prevented our recording the responding that occurred during these sessions.

As in Experiment 1b, there were two test sessions, each containing two presentations of X and Y. In other respects not specified here the procedure was the same as that described for Experiment 1b.

Results and discussion

All subjects initially acquired responding on the VI 60-s schedule, and the use of the lower intensity shock for reinforced trials allowed this responding to be maintained in all animals. PreCS and CS scores for each subject were pooled across all trials of a given type in Phase 1 in order to compute a single suppression ratio for each compound. The group means were .10 for AX and .05 for BY. No data were recorded during Phase 2.

In the previous experiments the test results were complicated by intrinsic differences in the ability of the stimuli used as X and Y to evoke suppression. Increasing the intensity of the light in the present experiment eliminated this problem—a preliminary inspection of the data revealed no overall difference between the light and dark cues in the suppression they evoked, and the results for the two cues are pooled in the data summarized in Figure 3. A further difference between this and the previous experiments was that X and Y evoked much less suppression during the test—indeed suppression was evident only on the first of the test trials. (This was presumably a consequence of the reduced Phase 1 shock intensity.) Accordingly, rather

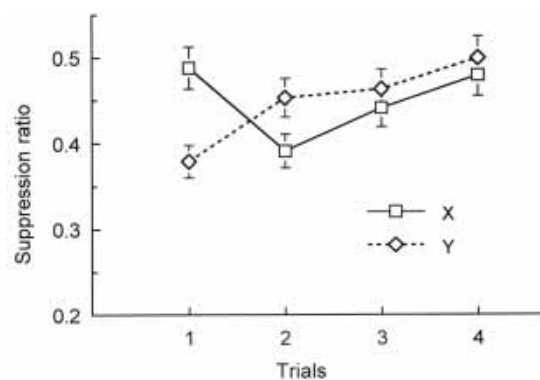


Figure 3. Experiment 1c. Group mean suppression scores for the test trials with X and Y. All animals had previously received reinforced trials with the compounds AX and BY followed by nonreinforced presentations of A. Error bars represent within-subjects standard error, computed on scores adjusted for variation between subjects.

TABLE 1
Experiment 1c: PreCS rates^a

PreCS	Trial			
	1	2	3	4
X	27.76	28.24	26.24	23.50
Y	28.86	27.74	23.50	23.36

^aResponse per min.

than pooling the data over all test trials we present, in Figure 4, trial-by-trial suppression ratios (see p. 344). The only sizeable effect was on Trial 1 where Stimulus Y evoked more suppression than Stimulus X, that is, the mediated extinction effect of the previous experiments was replicated on this trial. Thereafter, Stimulus X evoked slightly more suppression than Stimulus Y, although the differences were very small. An ANOVA with stimulus type and trial as variables was performed on these data. This revealed no main effect of stimulus type, $F < 1$, no main effect of trial, $F(3, 45) = 2.39$, but a significant interaction of these two factors, $F(3, 45) = 4.06$. The source of this interaction was examined with a test of simple main effects. This revealed a significant effect of stimulus on Trial 1, $F(1, 15) = 10.54$, but no significant effects on the other trials ($F_s < 2$). The preCS rates on which the ratio scores were based are shown in Table 1. These rates were higher than those in the previous experiments (perhaps because the Phase 2 treatment allowed extended experience of lever-press training), and although the rates tended to decline from one trial to the next, there is no indication of any other difference that might account for that seen in the suppression scores. An ANOVA conducted on the data summarized in Table 1, with trial type and trial number as the variables showed a significant effect only of the latter, $F(3, 45) = 2.85$ (other $F_s < 1$).

This experiment has again failed to demonstrate a recovery-from-overshadowing effect and allows the conclusion that using a large number of Phase 2 extinction trials is not in itself enough to generate the effect. Indeed the only significant effect obtained (on Trial 1 of the test) was in the opposite direction. Depending, as it does, on the effect seen on a single test trial, the result of this experiment does not constitute a particularly convincing demonstration of the mediated extinction effect; but, taken together with the results of Experiment 1a and 1b, it prompts the conclusion that the consequence of extinguishing the associate of a CS is, if anything, to reduce rather than enhance the magnitude of the CR governed by that stimulus. In all three experiments only a small effect was obtained, but in each it was in the direction of mediated extinction.

EXPERIMENT 2

Given the results of the three studies that constituted Experiment 1, we were unable to follow our original strategy, which required a clear demonstration of the recovery-from-overshadowing effect. Accordingly we decided on attempting to analyse the nature of the effect that we did obtain—by exploring the source of mediated extinction.

In the Introduction we outlined two possible explanations for this effect. Both of them were based on the assumption that within-compound associations will be formed during the initial

stage of reinforced training with the AX compound. According to the interpretation offered by Holland (1990) the existence of this association means that subsequent presentations of A will be able to activate the representation of X, with the result that X, as well as A, will undergo extinction during the nonreinforced A trials. For this analysis the association $A \rightarrow X$ plays the critical role. For the alternative, associative-chain account (Rescorla & Cunningham, 1978), on the other hand, it is the $X \rightarrow A$ association that is critical. According to this account, the CR governed by X will be determined, in part, by its ability to activate its associate. When the associate itself governs some associative strength, the CR will be augmented; but when, as in these experiments, the associate (A) has undergone extinction, the ability of X to activate the A representation will be irrelevant to the magnitude of the CR obtained.

This analysis suggests a way of distinguishing between the alternative explanations of the mediated extinction effect. In the procedure used in Experiment 1, A and X were presented simultaneously during Phase 1 of training, allowing for the formation of both $A \rightarrow X$ and $X \rightarrow A$ associations (or alternatively of a single bidirectional association between A and X). In the present experiment we adopted the same general experimental design, but modified the procedure by giving serial presentation of the cues in Phase 1, the offset of A immediately preceding the presentation of X. (The control stimuli, B and Y, were similarly presented in the serial $B \rightarrow Y$ arrangement.) This procedure should allow the formation of strong $A \rightarrow X$ and $B \rightarrow Y$ associations; according to standard associative theory, however, strong $X \rightarrow A$ and $Y \rightarrow B$ associations are unlikely to be established under these conditions. In the absence of the $X \rightarrow A$ and $Y \rightarrow B$ associations, the associative-chain process will be unable to operate—neither the target cue, X, nor the control cue, Y, will be able to activate its associate on test and, accordingly, the associative strength governed by the associate will be unable to influence test performance. Phase 2 presentation of A, therefore, should be without effect, and there should be no difference between X and Y in the CR they evoke. Holland's (1990) notion of mediated extinction, on the other hand, predicts that the difference should still be found; what is critical according to this account is that A should be able to activate the X representation during the Phase 2 trials, and this should still be possible when the rats are trained with the serial, $A \rightarrow X$, arrangement in Phase 1. A study using this experimental design in an appetitive conditioning procedure (Holland & Ross, 1981, Exp. 3) produced evidence of a mediated-extinction effect.

Method

Subjects and apparatus

The subjects were 16 experimentally naive male hooded (Lister) rats whose mean ad lib weight was 385 g (range: 330–410 g). The apparatus was that used in Experiment 1a, and the procedure was the same as that described for Experiment 1a apart from the exceptions specifically detailed below.

Procedure

After the instrumental baseline response had been established, the rats received 12 sessions of Phase 1 training. In an attempt to avoid the loss of baseline responding that occurred during Phase 1 training in the previous experiments, the number of reinforced trials was reduced to two per session. In each Phase 1 session the rats received two presentations of a serial compound followed by a 1.0-mA, 0.5-s footshock. For the first two sessions both trial types in a given session were the same. Half of the subjects received $A \rightarrow X \rightarrow US$ trials on Session 1 and $B \rightarrow Y \rightarrow US$ trials on Session 2; the remainder received the

opposite arrangement. In the remaining 10 sessions one trial of each type was delivered in each session. Half of the animals received the $A \rightarrow X \rightarrow US$ trial followed by the $B-Y-US$ trial on these sessions. The remaining animals were given the opposite arrangement. For half of the subjects A was the light, and B was the dark stimulus; for the remainder these stimulus arrangements were reversed. For half of the subjects in each of these groups, X was the clicker and Y the tone; for the remaining rats, the arrangement was the reverse. The duration of stimuli was 5 s for A , B , X , and Y in order to maintain the overall duration of the compound at 10 s, which was the duration used for the simultaneous compound in the experiments described previously. The ITI was 790 s. Responding was recorded during stimulus presentations and also during the 5-s stimulus-free period that immediately preceded presentation of the first element of the compound.

The procedure for Phase 2 was identical to that described for Experiment 1a except that the duration of the stimulus in the present experiment was 5 s. Thus in each of the next three sessions, all subjects received six nonreinforced presentations of A . Trials were separated by an ITI of 338 s. Sessions were 40 min in length. Responding was recorded during these trials and also during the 5-s preCS period.

Each of the two test sessions contained three presentations of X and three of Y . Half the animals experienced the trial sequence X,Y,Y,X,Y,X on the first of these sessions, and the sequence Y,X,X,Y,X,Y on the second; for the remaining animals the arrangement was reversed. As before, the duration of each stimulus presentation on test was 30 s. Responding was also recorded during the 30-s preCS periods. The ITI was 317 s.

Results and discussion

Baseline responding was well maintained during Phase 1, but suppression in the presence of the CSs was acquired very readily, being almost total in the presence of each of the cues after the second session of training. The nonreinforced training of Phase 2 resulted in a loss of suppression in the presence of Stimulus A . For each animal responding was pooled over all six trials on a given day, and a suppression ratio was calculated. The group mean scores were .14 on Day 1, .20 on Day 2, and .29 on Day 3. An ANOVA conducted on these data confirmed the reliability of the effect of day, $F(2, 30) = 6.23$.

For the test phase, suppression ratios were calculated from the total preCS and CS responding for each subject over all six trials with a given stimulus. The mean response rate in the preCS periods preceding trials with X was 35.13 responses per min; that in the preCS periods preceding trials with Y was 37.44 responses per min. These rates did not differ significantly ($F < 1$). The group mean suppression scores shown in Figure 4 indicate that suppression was more substantial in the presence of X than of Y ; that is, Phase 2 presentation of A enhanced the degree of suppression governed by X . This constitutes a recovery-from-over-shadowing effect. Statistical analysis supported this conclusion. An ANOVA conducted on the overall means presented in Figure 4 showed that the difference was significant, $F(1, 15) = 5.08$.

The results of this experiment were most unexpected. We used a procedure designed to distinguish between alternative interpretations of the mediated-extinction effect, specifically one designed to eliminate any contribution from the associative-chain process, described in the Introduction to the experiment. We anticipated that the effect would be abolished if it depended on the associative-chain process but would still be found if it depended on the mediated conditioning process postulated by Holland (1990). What we found was that the effect was not merely abolished, but was in fact reversed—that is, this procedure generated the effect

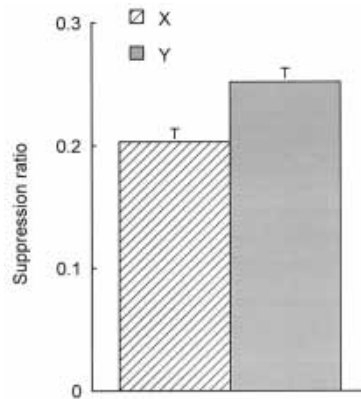


Figure 4. Experiment 2. Group mean suppression scores for Stimuli X and Y in the test phase. Error bars represent within-subjects standard error, computed on scores adjusted for variation between subjects.

that we had sought in Experiment 1, recovery from overshadowing. The implications of these results are taken up in the General Discussion.

GENERAL DISCUSSION

The three studies presented as Experiment 1 explored the claim that the recovery-from-overshadowing effect could be produced using a conditioned suppression procedure (e.g., Miller et al., 1992). In all three studies, rats received shock-reinforced trials with a simultaneous compound stimulus, AX, followed by nonreinforced presentations of A alone. We expected to find that this treatment would enhance the magnitude of the CR governed by X. No such effect was found, and in fact the results tended in the opposite direction in that a small mediated-extinction effect (i.e., an attenuation of the CR to X) was found in all cases. The recovery-from-overshadowing effect was, however, obtained in Experiment 2, which differed from the preceding experiments principally in the use of a serial (A → X) compound in the first stage of training.

In attempting to explain this pattern of results we begin with the mediated-extinction effect. As we noted in the Introduction to this article, there are currently two rival explanations for this phenomenon—the associative-chain account, which supposes that extinction of A reduces the CR by eliminating the contribution from the X → A association on test, and the mediated conditioning account, which supposes that the associatively activated representation of X itself undergoes extinction during nonreinforced A trials. Either of these processes could be responsible for the effects obtained in Experiment 1. In Experiment 2, however, we gave training with the serial A–X compound that was unlikely to establish the X → A association required by the associative-chain mechanism but could be expected to establish the A → X association required by the mediated-conditioning account. The results provided no evidence of a mediated-extinction effect and thus no evidence for the proposed mediated-conditioning mechanism. Although the absence of a mediated-conditioning effect in Experiment 2 cannot prove that the effect does not operate in the training conditions employed in

Experiment 1, the most parsimonious conclusion is that the results obtained in that experiment were produced by the associative-chain mechanism.

What remains is to explain why Experiment 2 should have yielded the reverse of mediated extinction—an enhancement of the CR to Stimulus X. It seems that when the training conditions are such as to preclude operation of the associative-chain mechanism, the effects of some other learning process are able to show themselves. We conclude, therefore, that one or other of the mechanisms proposed for retrospective revaluation is operating in these experiments. It is certainly possible that, for the procedure used in Experiment 1, presentation of A will both reduce the effectiveness of the A-US association (producing mediated extinction via the associative-chain mechanism) and also allow the acquisition of associative strength by the associatively activated representation of X, as postulated by Dickinson and Burke (1996). It may also be possible for comparator theory to incorporate the associative-chain mechanism—to allow that the value of Stimulus A might affect test performance to Stimulus X both by way of its own strength (i.e., by way of the X-A chain) and also by way of its comparator function. In either case the outcome of extinction of A following reinforced training with the simultaneous AX compound would depend on the balance of two processes, one tending to attenuate the CR governed by Stimulus X and the other to increase it. In our Experiment 1 the first of these processes appears to dominate, so that the effect of the other only becomes evident when, as in Experiment 2, this process is not able to operate.

Although the two-process account just outlined provides a coherent explanation for the experimental results reported here, problems arise in applying it more generally. First, the mechanism we have endorsed for mediated extinction implies that the effect should not be found when animals receive serial, $A \rightarrow X$, training in Phase 1. But the demonstration of mediated extinction reported by Ward-Robinson and Hall (1996) used just such a procedure. It should be noted, however, that the procedure used by Ward-Robinson and Hall differed in a number of ways from that used in the present Experiment 2. First, stimulus durations were rather different (A was longer, and X was very brief in the initial phase of training). Second, the Phase 1 treatment of the present Experiment 2 was divided up into two separate stages in the Ward-Robinson and Hall experiment; that is, the rats received a set of nonreinforced $A \rightarrow X$ trials before being given reinforced training with X as a separate block of trials. It is not clear why these procedural differences should generate such a different outcome, but experiments to isolate the relevant factor can readily be devised. Less tractable is the second problem raised by our results. Our demonstrations of the mediated-extinction effect in Experiment 1 made use of a simultaneous AX compound in the first phase of training and led us to conclude that this effect would be likely to obscure any recovery-from-overshadowing effect under these training conditions. But, as we have already noted, Miller et al. (1992) have successfully obtained the latter effect using just such a training procedure. We must assume that there is some other factor at work in the experiments by Miller and his colleagues that tips the balance in favour of recovery from overshadowing, but we have no notion at this stage as to what this factor might be.

Although problems remain, it is appropriate to conclude by emphasizing the positive contribution made by the experiments reported here. First, they demonstrate that both the mediated-extinction effect and the recovery-from-overshadowing effect can be obtained in a standard conditioned suppression paradigm. The former had previously been shown only

in a study using rather unusual parameters (Ward-Robinson & Hall, 1996); the latter (apart from the early study by Kaufman & Bolles, 1981) only in the specific training preparation used by Miller and his colleagues (Miller et al., 1992). Second, the interpretation we have offered for our results provides an explanation for why it should often have proved difficult to obtain reliable demonstrations of either of these effects. If it is accepted that two processes are at work in this procedure—one (which we have identified as the associative-chain mechanism) tending to produce mediated extinction, and another (yet to be decisively identified) that tends to produce recovery from overshadowing—then the outcome of any given experiment will depend on those factors that determine the relative effectiveness of these two processes. We do not yet know what these factors might be, but it is easy to imagine that they will differ according to the specific training procedure used, allowing both effects to emerge under certain circumstances and, when the balance between the processes is even, sometimes resulting in no effect at all.

REFERENCES

- Bakeman, R., & McArthur, D. (1996). Picturing repeated measures: Comments on Loftus, Morrison, and others. *Behavior Research Methods, Instruments, & Computers*, *28*, 584–589.
- Blaisdell, A. P., Bristol, A. S., Gunther, L. M., & Miller, R. R. (1998). Overshadowing and latent inhibition counteract each other: Support for the comparator hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*, 335–351.
- Blaser, R. E., Couvillon, P. A., & Bitterman, M. E. (2004). Backward blocking in honeybees. *Quarterly Journal of Experimental Psychology*, *57B*, 349–360.
- Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective reevaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, *49B*, 60–80.
- Hall, G. (1996). Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning & Behavior*, *24*, 233–255.
- Holland, P. C. (1981). Acquisition of representation-mediated conditioned food aversion. *Learning and Motivation*, *12*, 1–12.
- Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition*, *37*, 105–131.
- Holland, P. C. (1999). Overshadowing and blocking as acquisition deficits: No recovery after extinction of overshadowing or blocking cues. *Quarterly Journal of Experimental Psychology*, *52B*, 307–333.
- Holland, P. C., & Forbes, D. T. (1982). Representation mediated extinction of conditioned flavour aversions. *Learning and Motivation*, *13*, 454–471.
- Holland, P. C., & Ross, R. T. (1981). Within-compound associations in serial compound conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*, 228–241.
- Kaufman, M. A., & Bolles, R. C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society*, *18*, 318–320.
- Konorski, J. (1967). *Integrative activity of the brain*. Chicago: University of Chicago Press.
- Larkin, M. J. W., Aitken, M. R. F., & Dickinson, A. (1998). Retrospective reevaluation under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1331–1352.
- Matzel, L. D., Shuster, K., & Miller, R. R. (1987). Covariation in conditioned response strength between elements trained in compound. *Animal Learning & Behavior*, *15*, 439–477.
- Miller, R. R., Barnet, R. C., & Graham, N. J. (1992). Responding to a conditioned stimulus depends on the current associative status of the cues present during training of that specific stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*, 251–264.
- Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 51–92). San Diego, CA: Academic Press.

- Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, *94*, 61–75.
- Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, *101*, 587–607.
- Rescorla, R. A., & Cunningham, C. L. (1978). Within-compound flavour associations. *Journal of Experimental Psychology: Animal Behavior Processes*, *4*, 267–275.
- Rescorla, R. A., & Wanner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, *25*, 127–151.
- Wagner, A. R., & Brandon, S. E. (2001). A componential theory of Pavlovian conditioning. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 23–64). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Ward-Robinson, J., & Hall, G. (1996). Backward sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*, 395–404.

Original manuscript received 23 June 2003
Accepted revision received 11 September 2003

