Learned irrelevance: A contemporary overview

Charlotte Bonardi and Siaw Yann Ong

University of York, York, UK

This article reviews the recent literature on the topic of learned irrelevance. It asks whether the retardation of subsequent conditioning produced by uncorrelated preexposure is indeed the result of the animal learning that a conditioned stimulus (CS) and unconditioned stimulus (US) are unrelated, or whether it is better explained either as a result of the context specificity of latent inhibition, or as some other artefact of the uncorrelated schedule employed. The conclusion is that there is as yet no good evidence to support the existence of a “genuine” learned irrelevance effect.

Associative theory originally assumed that the associability of a stimulus was directly related to its salience and did not change with experience (e.g., Rescorla & Wagner, 1972; cf. Mackintosh, 1975). Moreover, there were no particular constraints on what could be associated with what—one stimulus was as able to become associated with one outcome as with another. But it soon became apparent that these assumptions were too simplistic. The phenomenon of latent inhibition demonstrated that nonreinforced preexposure to a stimulus could severely retard the degree to which it could enter into associations (Lubow & Moore, 1959)—a finding that had to be incorporated into associative theorizing. Moreover, one could also produce a type of latent inhibition by preexposing the unconditioned stimulus (US)—the “US preexposure effect” (e.g., Kremer, 1971). It appeared that there were multiple mechanisms that could affect the degree to which one stimulus could become associated with another.

Latent inhibition and the US preexposure effect are both nonspecific phenomena—a latently inhibited stimulus shows impairments in becoming associated with any outcome, and preexposing the US will impair its ability to become associated with any CS. But there are more subtle possibilities. Baker and Mackintosh (1977; also, e.g., Baker, 1976) reported that when two stimuli are presented in an uncorrelated manner, such that one bears no predictive relation to the occurrence of the other, subsequent formation of excitatory or inhibitory associations between those two stimuli is retarded. This phenomenon they termed learned...
irrelevance, and Baker (1976) interpreted it as evidence that the animal had learned about the absence of correlation between those two specific events.

Demonstrating an effect of this type, however, has been dogged by experimental problems. First, uncorrelated preexposure requires presentations of both the conditioned stimulus (CS) and the US. But CS exposure produces latent inhibition, and US presentation produces the US preexposure effect; if learned irrelevance exists, it should create a retardation of learning greater than that produced by the sum of CS and US preexposure (Baker & Mackintosh, 1977). Although some studies have attempted to demonstrate retardation that conforms to these criteria (Baker & Mackintosh, 1979; Matzel, Schachtman, & Miller, 1988), interpretative problems remain (see Baker, Murphy, & Mehta, 2003). Moreover, the strategy these studies employed was to examine whether manipulations known to reduce latent inhibition and the US preexposure effect would eliminate the retardation of learning produced by uncorrelated preexposure. As some retardation of learning survived these manipulations, the authors of these studies argued that learned irrelevance was not the sum of these two effects. However, this approach is logically problematic unless the critical manipulations eliminate both latent inhibition and the US preexposure effect; if they do not, then any residual “learned irrelevance” might really reflect only the sum of two subthreshold effects—latent inhibition and US preexposure—which when added together are capable of affecting behaviour. And as there is no guarantee that either latent inhibition or the US preexposure effect were eliminated, as opposed to merely reduced, the results of these studies must be taken only as tentative evidence for the existence of learned irrelevance.

However, more recent work by Mackintosh and his colleagues (Bennett, Maldonado, & Mackintosh, 1995) adopted a different approach to deal with the problem of whether learned irrelevance is no more than the sum of latent inhibition and US preexposure. They used the technique of preexposing one group of animals to an uncorrelated schedule of CS and US presentations, and another to the same number of CSs and USs, but in separate sessions. As both groups receive the same number of CS and US presentations during preexposure, the retardation of subsequent learning due to latent inhibition and the US preexposure effect should be the same in both. But as only the first group has the opportunity to learn that CS and USs are uncorrelated, only this group should show learned irrelevance. Using this approach they demonstrated that uncorrelated preexposure could produce a retardation in the acquisition of excitatory conditioning that could not be explained as the sum of latent inhibition and the US preexposure effect.

Context specificity of latent inhibition

Although this approach neatly sidestepped the procedural problems that had dogged prior attempts to demonstrate learned irrelevance, it introduced others. It has been noted (Baker & Mackintosh, 1979; Bonardi & Hall, 1996) that another factor that may confound interpretation of the results of experiments on uncorrelated preexposure is the context specificity of latent inhibition (e.g., Channell & Hall, 1983; Lovibond, Preston, & Mackintosh, 1984). If a stimulus is preexposed in a distinctive context, then the latent inhibition that would normally be seen to have accrued to that stimulus may be attenuated if conditioning takes place in a different context. Moreover, the type of cue that may produce such contextual dependence includes the presence of a US (Bouton, Rosengard, Achenbach, Peck, & Brooks, 1993;
The problem that arises from this observation is that in the procedure employed by Mackintosh and his colleagues (Bennett et al., 1995), the uncorrelated schedule involved preexposing the CS in the presence of the US and its aftereffects, whereas the control procedure did not. As in the conditioning phase the animals received several trials per session, the majority of conditioning trials also occurred in the presence of US aftereffects. Thus animals receiving uncorrelated preexposure would have experienced both preexposure and conditioning in the same context, that of US presentations, and should therefore show strong latent inhibition. The control animals, in contrast, would have been preexposed to the CS in the absence of US aftereffects and conditioned in their presence. They would therefore suffer a change of context between preexposure and conditioning that would attenuate the latent inhibition produced by preexposure, and so the rate of conditioning in these animals should be enhanced. Thus the results observed by Bennett et al. (1995)—slower excitatory conditioning after uncorrelated preexposure—may be explained completely in terms of the context specificity of latent inhibition, without appealing to the notion of learned irrelevance.

Bonardi and Hall (1996) tested this idea. They reasoned that if it were possible to condition animals in the absence of US aftereffects, then the opposite pattern of results should be observed: Now the animals given uncorrelated preexposure would be the ones to suffer a context change between the preexposure and conditioning phases, and they should therefore be the ones to show an attenuated latent inhibition effect. In contrast, the control animals, being both preexposed and conditioned in the absence of US aftereffects, should show a robust latent inhibition effect—and hence slower learning.

But how to condition the animals in the absence of US aftereffects? Bonardi and Hall (1996) attempted to achieve this by giving animals one conditioning trial per session. Thus in a conditioned suppression procedure, one group of animals received uncorrelated preexposure to a noise CS and shock US, whereas control subjects received a block of sessions in which they were preexposed to the noise, followed by a block of sessions in which they were preexposed to the shock. Then all subjects received pairings of the noise and shock, with one conditioning trial per session. The results of this experiment are shown in Figure 1, where it may be seen that, in accord with Bonardi and Hall’s predictions, uncorrelated preexposure produced significantly faster conditioning than preexposure in which CS and US were presented in separate blocks of sessions. As they also succeeded in demonstrating that animals given several conditioning trials per session in the test conditioned more slowly after uncorrelated preexposure, they argued that their procedures were sufficient to produce a normal “learned irrelevance” effect, but that this effect was best explained in terms of the context specificity of latent inhibition, rather than to the animals learning that CS and US were unrelated.

Unsurprisingly, Mackintosh and his colleagues challenged this interpretation. In a subsequent article (Bennett, Wills, Oakeshott, & Mackintosh, 2000) they reasoned that if the effects of uncorrelated preexposure were entirely the product of the context specificity of latent inhibition, then certain testable predictions would follow. First, they argued that in a multitrial conditioning procedure, animals given uncorrelated preexposure should show more conditioning on the first trial of each session than control subjects, as this trial occurs in the absence of the aftereffects of the US. They investigated this and found the opposite result—animals given uncorrelated preexposure actually learned less than control subjects on this first trial. Second, they conducted an experiment in which, after the standard preexposure treatments, all subjects received one conditioning trial per session, and this trial was always...
preceded by an excitatory clicker; for half the subjects this clicker was reinforced, whereas for the remainder it was nonreinforced. They argued that if the presence of US aftereffects is critical for the demonstration of learned irrelevance, then animals for whom the clicker was reinforced should show learned irrelevance and hence slower conditioning after uncorrelated preexposure. Conversely animals for whom the clicker was nonreinforced should show the opposite effect—faster conditioning after uncorrelated preexposure. But this was not what was observed: Uncorrelated preexposure produced slower conditioning regardless of whether the clicker that signalled the conditioning trials was reinforced or nonreinforced. In a final experiment they attempted to replicate Bonardi and Hall’s (1996) finding of faster conditioning after uncorrelated preexposure when animals receive only one trial per session in the conditioning phase; once more they failed. Uncorrelated preexposure again retarded excitatory conditioning.

As a result of these observations, Bennett et al. (2000) concluded that their original interpretation (Bennett et al., 1995) had not been successfully challenged. They argued that the shock US used by Bonardi and Hall (1996) would probably produce aftereffects that were substantially stronger than those of the sucrose reward used in the appetitive procedures that they had employed. If the US aftereffects are sufficiently strong, a context that is able to modulate the expression of latent inhibition would be produced. Thus in the one-trial conditioning procedure a profound loss of latent inhibition would be observed in animals given uncorrelated preexposure, and this might be enough to obscure any learned irrelevance effect that was also present. But if these US aftereffects were to be reduced, as in Bennett et al.’s (2000) experiments, this context specificity of latent inhibition would be attenuated, and the learned irrelevance effect would be revealed. There was no need, therefore, as yet to doubt the existence of learned irrelevance.
The nature of the preexposure schedule

The data presented by Bennett et al. (2000) do cast doubt on the suggestion that US aftereffects are playing a significant role in their appetitive procedure and hence that the context specificity of latent inhibition is a complete explanation of the learned irrelevance effect. But there are further potential problems. It is incumbent on anyone attempting to demonstrate learned irrelevance to ensure that the uncorrelated schedule that they employ is not producing a CS with inhibitory properties—as inhibitory conditioning in itself could, of course, produce a retardation of excitatory learning. Conscious of this requirement, Bennett et al. (2000) investigated this issue. In Experiment 1b of the Bennett et al. article one group of animals received an uncorrelated preexposure treatment in which, on each of the 10 days of the preexposure phase, they received one session of uncorrelated CS and US presentations. Control animals received the same number of CS and US presentations, but in separate sessions—they received all their US presentations in the first half of the preexposure phase and all their CS presentations in the second. Then all subjects received training in which the preexposed CS signalled nonreinforcement of an excitatory clicker—training designed to establish the preexposed CS as a conditioned inhibitor. Inhibitory conditioning was acquired more slowly by animals given uncorrelated preexposure, which is consistent with their proposal that their preexposure schedule was producing learned irrelevance, rather than inhibitory conditioning.

But a sceptic might argue that this result is not enough to allow us to dismiss the suggestion that the uncorrelated schedule was producing conditioning inhibition. This is because in this experiment, as in all the others in the Bennett et al. (2000) article, the uncorrelated group received one session of uncorrelated CS and US presentations on every day of preexposure, whereas control subjects received all their USs in the first half of this phase, and only CSs in the second. This design raises the possibility that the context–US association at the start of the test might differ in the two groups, because during the second half of preexposure it would have the opportunity to extinguish for subjects in the control, but not in the uncorrelated, group. A difference in context conditioning between the two groups at the start of conditioning could at least have affected the results and at worst have masked the inhibitory effects of the uncorrelated schedule. Given the theoretical importance of this demonstration, it seemed important to rule out this alternative interpretation of the results, and so the issue was explored in a study performed in this laboratory, in collaboration with Siaw Yann Ong and Geoffrey Hall (Bonardi, Hall, & Ong, 2002).

The question of interest was whether the difference in the distribution of USs between two groups, the first of which received a session with USs on each day of preexposure and the second of which received all their USs in the first half of this phase, could in itself affect the subsequent acquisition of inhibitory conditioning. Accordingly two groups of animals received preexposure identical to that in the Bennett et al. (2000) study, except that all CS presentations were omitted during the preexposure phase. After these treatments animals in the two groups received inhibitory conditioning to the target CS—the 10-s presentation of two jewel lights. The inhibitory conditioning sessions were like those described in Bennett et al. (2000, Exp. 1b), except that on compound trials the light was presented immediately prior to the 10-s test excitor, rather than in a simultaneous compound with it, and that there were four reinforced and two nonreinforced trials in each session. The test excitor, a 10-Hz 75-dB
clicker, was established in exactly the same way as that in Bennett et al. (2000, Exp. 1b), except that in this study animals received a total of three rather than four clicker–food trials per session; the reinforcer was the delivery of four food pellets. The results of this experiment are shown in Figure 2. The measure of conditioned inhibition was a ratio score of form a/a+b, where a was the rate of responding during the clicker when it was signalled by the light, and b the rate during the clicker when it was presented alone. When the light has no inhibitory properties, and hence no effect on responding to the clicker, a and b will be equal, and the ratio .5. The more inhibitory power is acquired by the light, the lower a will become relative to b, and hence the lower the resultant ratio. It is evident in the figure that conditioned inhibition was acquired more slowly in the animals that had received the same pattern of US delivery as those given uncorrelated preexposure in Bennett et al.’s (2000) experiment. This conclusion is supported by the results of an analysis of variance with group and sessions as factors, which revealed a significant Group × Sessions interaction, $F(2, 26) = 3.5$. (The criterion for statistical significance throughout this article is that $p < .05$.) Analysis of simple main effects revealed that there was a significant difference between the two groups on Session 2, $F(1, 36) = 4.79$. These results clearly indicate that the difference in the pattern of US distribution in the two groups during preexposure had a marked effect on inhibitory conditioning. (The results of a companion experiment that examined the effect of these two preexposure treatments on the course of subsequent excitatory conditioning to that CS, using the excitatory conditioning procedure employed by Bennett et al., 1995, found no effect; an analysis of variance with group and sessions as factors revealed no effect of group or Group × Sessions interaction, $F$s < 1.)

![Figure 2](image)

**Figure 2.** Group mean suppression ratios (a/a+b, where a is the mean response rate on reinforced clicker trials, and b the mean rate on nonreinforced clicker trials) for the three test sessions for group U (US distribution equivalent to the group given uncorrelated preexposure in Bennett et al., 2000, Exp. 1b) and group C (US distribution equivalent to the control group of the same experiment). The bars show standard errors.
Learned irrelevance or conditioned inhibition?

The finding described in the preceding paragraph was of concern, as it suggested that the difference in the distribution of US presentations between the two groups in Bennett et al.’s (2000) study was in itself sufficient to retard the acquisition of conditioned inhibition. This difference could, therefore, be responsible for the retardation of inhibitory conditioning observed by Bennett et al. (2000)—meaning that the true effect of their uncorrelated schedule on inhibitory learning was as yet unclear. The next step was to examine the effect of their uncorrelated schedule in the absence of this confound in the distribution of USs during the preexposure phase. To investigate this issue further, two more experiments were performed. In both experiments one group of animals received one session of uncorrelated preexposure to CSs and USs on each day of preexposure, exactly as in the Bennett et al. (2000) experiments; but to equate the pattern of US preexposure between this experimental group and the control group, control subjects also received US presentations on every day of preexposure; this matching was achieved by giving them one session of CSs and another of USs on each day of preexposure. The distribution of CSs and USs within these sessions was otherwise the same as that in the study by Bennett et al. (2000). Then for some subjects the CS was trained as a conditioned excitor and for others as a conditioned inhibitor. The test procedures were identical to those employed in the previous two experiments, except that the test excitor training sessions consisted of four rather than three trials. The results are shown in Figures 3 and 4 and are quite clear. First, uncorrelated preexposure retarded the acquisition of conditioned excitation (Figure 3): An analysis of variance with group and sessions as factors revealed a significant Group × Sessions interaction, $F(5, 145) = 2.71$, and simple main effects analysis revealed an emerging trend in that the groups differed almost significantly on Session 4, $F(1, 97) = 3.84$, $p = .053$, and significantly on Session 6, $F(1, 97) = 8.01$. However, uncorrelated preexposure
facilitated the acquisition of conditioned inhibition (Figure 4); an analysis of variance with group and two-session blocks as factors revealed a significant Group × Blocks interaction, $F(3, 90) = 3.83$; simple main effects analysis revealed that the groups differed significantly on Block 2, $F(1, 105) = 12.14$.

After removing the confound in the distribution of USs during the preexposure phase, it appears that the uncorrelated preexposure schedule employed in the Bennett et al. (2000) studies is actually producing a CS with inhibitory properties. Moreover, these inhibitory properties must be substantial, given that the treatment given to the control groups of these two experiments—alternating sessions of CS and US exposure—can itself produce demonstrable conditioned inhibition in the CS (Baker, 1977). The reason for this is unclear, although one possibility lies in the details of the schedule. In the uncorrelated preexposure sessions, CS presentations occurred with an average intertrial interval (ITI) of 2 min (range 1–3 min). Each ITI was divided into 10 equal time bins, 1–10. Bin 1 immediately followed the offset of the CS, and Bin 10 immediately preceded the 10-s preCS period that immediately preceded each CS presentation. Within each session, one US was randomly presented once at the start of each of Bins 1–10, with the following restrictions: No more than two USs could occur during any one ITI, and the presentation of two USs during a single ITI could only occur once per session. The requirement that only one ITI in any session could contain more than one US results in a relatively predictable alternation between CSs and USs that could conceivably result in the CS predicting the absence of the US—and preliminary work in this laboratory indeed suggests that when this alternation is eliminated, the apparent inhibitory effect disappears. But whatever the cause, the fact remains that this schedule produces a CS that retards excitatory conditioning but facilitates inhibitory conditioning. The implication is that the uncorrelated schedule is producing inhibitory conditioning to the preexposed CS that could have been responsible for the ability of uncorrelated
preexposure to retard excitatory conditioning in both the one- and the multitrial case in the studies by Bennett et al. (2000).

Conclusions

In conclusion, the evidence for the existence of learned irrelevance effect has proved equivocal. The results of those experiments that show it to be more than the sum of latent inhibition and the US preexposure effect can be explained in other ways—the aversive experiments (Bonardi & Hall, 1996) in terms of the context specificity of latent inhibition, and the appetitive experiments (Bennett et al., 1995, 2000) in terms of the specific properties of the schedule that was employed. It seems that there is as yet little conclusive evidence to suggest that an animal may learn that one event is unrelated to another.

Some authors (e.g., Baker, Murphy, & Mehta, 2003) have argued that it is theoretically suspect to explain the effects of uncorrelated preexposure in terms of two other effects—latent inhibition and the US preexposure effect—that are themselves lacking a firm theoretical explanation. But while the origins of latent inhibition and the US preexposure effect may be unclear, it is perhaps fair to say that the effects of preexposure to CS alone, and to US alone, on subsequent conditioning are well documented and robust. At an empirical level, therefore, it seems parsimonious at least to attempt to account for the effects of uncorrelated preexposure in terms of these two, simpler procedural manipulations—although, as we have seen, several other factors may explain the retarded conditioning seen after uncorrelated preexposure.

One final comment concerns whether it would ever be possible to demonstrate satisfactorily that uncorrelated preexposure produces learned irrelevance. As this article has suggested, to provide such a demonstration a number of criteria must be fulfilled. First, any control condition must involve preexposure to the same number of CSs and USs as occur in the uncorrelated condition; one such control is the between-session exposure to CSs and USs that was employed in the final two experiments of the present article. This control has the advantage of approximately equating between the experimental and control conditions the distribution of USs in the exposure phase—a factor that we have seen can have a profound effect on the results obtained. But there is a problem with this control. To ensure that the uncorrelated schedule is “truly” uncorrelated, it must be shown to produce a retardation of both excitatory and inhibitory conditioning; however, it has been argued that alternating preexposure to both CSs and USs used in the between-sessions control can in itself produce conditioned inhibition (Baker, 1977). If animals given uncorrelated preexposure develop conditioned inhibition more slowly than such a control, this may not reflect learned irrelevance in these animals, but rather conditioned inhibition in the control subjects. Thus an additional control is required that cannot result in conditioned inhibition. One possibility is a group of animals who receive all their CS presentations in the first half of preexposure and all their US presentations in the second; here, however, the matching of the US distribution between the experimental and control conditions is lost. Finally, the possibility of explaining the results in terms of the context specificity of latent inhibition must be ruled out—by, for example, also demonstrating a retardation of excitatory conditioning with only one conditioning trial per session. It should be clear that satisfying all these conditions is a daunting prospect. It may be a number of years yet, therefore, before evidence for a “genuine” learned irrelevance effect is secured.
REFERENCES


