

Backward Conditioning: A Reevaluation of the Empirical Evidence

Marcia L. Spetch, Donald M. Wilkie, and J. P. J. Pinel
University of British Columbia, Vancouver, British Columbia, Canada

There is an apparent discrepancy between the widespread view that backward conditioning does not occur and the experimental evidence which suggests that it does. Backward pairing of conditioned and unconditioned stimuli frequently has resulted in effects similar to those produced by forward pairing, and the results of several recent experiments have established that such effects cannot be attributed to factors other than stimulus pairing per se. Surprisingly, even some of the earlier experiments that provided the basis for the current skepticism concerning backward conditioning provide evidence of its existence. The failure to recognize backward conditioning as a legitimate phenomenon seems to reflect theoretical biases rather than a paucity of empirical evidence. Thus backward conditioning and its properties merit renewed interest and reexamination.

In a recent experiment by Keith-Lucas and Guttman (1975), rats that received a single footshock followed by the presentation of a toy hedgehog subsequently avoided the hedgehog. Such a finding probably would not surprise one unfamiliar with the animal learning literature, although the imagination of the experimenters might. As both Hudson (1950) and Keith-Lucas and Guttman pointed out, common sense would lead one to expect animals to have the ability to respond defensively to a novel stimulus detected after a sudden aversive event. An animal that sighted an unfamiliar predator following an abortive attack surely would not submit to another attack (i.e., a forward pairing of the predator and pain) before reacting defensively.

It is therefore interesting that the occurrence of such backward conditioning would not be predicted from a survey of the animal learning literature. The prevailing view is that conditioned responses (CRs) develop as a result of stimulus pairing only when the conditioned stimulus (CS) precedes the un-

conditioned stimulus (US). In most textbook discussions of classical conditioning (e.g., Hulse, Deese, & Egerth, 1975; Kimble, 1961; Millenson & Leslie, 1979; Osgood, 1953; Schwartz, 1978), the unqualified conclusion is that conditioning does not occur if the CS follows the US (but cf. Fantino & Logan, 1979). Opinions on what does occur as a result of such backward pairings, however, are divided. Mackintosh (1974) and Terrace (1973), among others, believe that backward pairings have little effect and that the occasional instances of apparent CRs can be attributed to factors such as pseudoconditioning or sensitization rather than to stimulus pairing. The wide acceptance of this view is illustrated by the fact that backward pairings have been frequently used as a baseline control in evaluating the conditioning resulting from forward pairings (Kalish, 1954; Moeller, 1954; Spence & Runquist, 1958). In contrast, Moscovitch and LoLordo (1968), for example, suggested that backward pairings result in conditioned inhibition (i.e., a behavior change opposite to that produced by forward pairings).

What is the empirical basis for the belief that the outcomes of forward and backward pairing differ? Pavlov (1927/1960, p. 27) reported the first failure to demonstrate backward conditioning, and his conclusion that backward conditioning does not occur has been defended in subsequent reviews of

This research was supported in part by a scholarship and grants from the Natural Sciences and Engineering Research Council of Canada. S. Gray, P. Pasquali, and F. Valle provided many valuable comments.

Requests for reprints should be sent to Marcia L. Spetch, University of British Columbia, Department of Psychology, 2075 Westbrook Mall, Vancouver, British Columbia, Canada V6T 1Y7.

the literature. Cautela (1965) reviewed several American backward conditioning studies and concluded, "In all of the preceding studies backward conditioning was not obtained or was attributed to pseudoconditioning (by the investigators themselves or by others)" (p. 140). In perhaps the most authoritative of contemporary surveys of the conditioning literature, Mackintosh (1974) concluded, "All in all there is little reason to accept the reality of backward conditioning" (p. 60). Only one major review provides an exception to this general conclusion. After reviewing evidence from several Soviet and American experiments, Razran (1956) asserted that backward conditioning produced genuine CRs and was not a case of pseudoconditioning. His review was either largely ignored or criticized (e.g., Mackintosh, 1974), however, and has had little impact on the current status of backward conditioning within the American learning literature.

Why were Keith-Lucas and Guttman (1975) able to demonstrate a phenomenon that has apparently eluded other investigators for so many years? Keith-Lucas and Guttman attributed their success to the nature of their CS. The hedgehog CS of their experiment was a complex stimulus that perhaps resembled a natural predator of the rat, whereas the CSs in previous attempts to demonstrate backward conditioning typically consisted of arbitrary stimuli such as lights and tones. They suggested that the backward conditioning established in their experiment reflected "an evolved capability of the rat and many similar species to associate certain classes of stimulus objects with exteroceptive pain" (p. 475).

Although such an interpretation meshes well with the recent interest in adaptive specializations in learning (e.g. Rozin & Kalat, 1971) and may account for some of the failures of previous investigators to demonstrate backward conditioning, other recent experiments (Heth & Rescorla, 1973; Mahoney & Ayres, 1976; Wagner & Terry, 1975) have successfully demonstrated backward conditioning with traditional CSs such as lights and tones. Thus a closer examination of the discrepancy between these recent

demonstrations of backward conditioning and the widespread belief that such conditioning does not occur is warranted.

Empirical Requirements for a Demonstration of Backward Conditioning

Is there an unequivocal demonstration of backward conditioning? The answer obviously depends on the particular criteria used as a basis for the evaluation. In evaluating whether a phenomenon exists, the criteria should be strictly empirical ones. Empirical criteria require that the phenomenon can be demonstrated and can be attributed to the experimental manipulation in question (in this case to the backward pairings of a US and a CS). Theories of conditioning often incorporate additional criteria that may be used to evaluate whether the phenomenon supports the theory. But such theoretical criteria should not be used to evaluate evidence for the existence of a phenomenon because the same evidence may be evaluated differently: positively using the criteria of one theory, negatively using criteria of a different theory. Four empirical requirements are basic to a demonstration of backward conditioning.

1. The most obvious requirement is, of course, the backward pairing of the US and CS. There is considerable disagreement in the literature as to what constitutes backward pairing. In some classification schemes (e.g., Kimble, 1961), all procedures in which CS onset follows US onset are defined as backward pairing. In this article we reserve the term *backward pairings* for only the most unambiguous of such cases—ones with no temporal overlap between the US and the CS. Evidence from only those experiments in which CS onset follows US offset (immediately or after a delay) are considered. According to traditional views, however, none of the procedures in which CS onset follows or even occurs simultaneously with US onset would be expected to produce conditioning.

There have been several recent reports of backward conditioning of taste aversions (cf. Logue, 1979). Although these may represent genuine instances of backward conditioning,

the problem of specifying the time course of the US (toxicosis) in the taste aversion paradigm makes it difficult to determine whether a true backward pairing has been used. For this reason, we do not evaluate evidence for backward conditioning of taste aversions.

2. Evidence must be provided that the observed behavioral change is a direct consequence of the backward pairing of the US and the CS. Thus the experiments must include controls that enable the investigator to specify the degree to which behavioral changes resembling CRs may have resulted from exposure to the US or CS rather than to their pairing. Two processes that could be confused with backward conditioning in inappropriately controlled studies have been identified. *Pseudoconditioning* is a change in behavior to a CS that results from prior exposure to the US (Grether, 1938). For example, an animal that has just received a painful shock may exhibit a fear response to a subsequent novel stimulus. If this occurs, regardless of whether this stimulus has been paired with shock, the fear response can be attributed to pseudoconditioning. *Sensitization* is an augmentation of an unconditioned response (UR) to a CS as a result of prior exposure to the US (Grant & Adams, 1944) or to the CS (Groves & Thompson, 1970). Sensitization controls are particularly important whenever the naturally occurring responses (i.e., URs) to the CS topographically resemble the CRs that are established through pairings with the US.

Random occurrence of both the CS and US (Rescorla's 1967, "truly random control procedure") generally is held to be the most appropriate control procedure for both pseudoconditioning and sensitization. This procedure is impractical, however, when conditioning occurs in a few trials. For example, with only two random presentations of the CS and US, it is possible to end up with two forward or two backward pairings. For studies of backward conditioning in which conditioning is asymptotic in a few trials, the best control procedure is one in which the CS and US are *explicitly unpaired* (i.e., the CS and US are both presented, but never close in time). Discriminative conditioning in which the US is paired with one CS but

not another is a version of this explicitly unpaired control (cf. Rescorla, 1967).

The explicitly unpaired procedure is preferable to either the CS-alone or US-alone procedures because it controls for effects of prior exposure both to the CS and to the US. Rescorla (1967) argued that the explicitly unpaired procedure introduces a negative CS-US contingency that may endow the CS with inhibitory properties. He also argued, however, that the same negative CS-US contingency is present in the backward procedure. Since any inhibitory tendencies that might result from such a contingency should be reflected equally in both groups, the unpaired procedure seems to be an appropriate control for backward conditioning.

3. The interval between trials must be long enough to rule out the possibility that apparent backward conditioning effects are the result of delayed forward pairings of the CS presented on one trial and the US presented on the following one. Alternatively unpaired control groups that are exposed to the same intervals between the CS and the next US presentation also can be used to assess the possible effects of such delayed forward pairing. Of course, the most effective way to completely rule out delayed forward pairing as a factor in backward conditioning is to limit conditioning to a single US-CS pairing.

4. To rule out inhibitory conditioning, backward pairing of a US and a CS must result in a behavioral change to the CS that is qualitatively similar to that produced by forward pairing, although the magnitude of this change and its temporal course need not be the same as that produced by a forward pairing. Furthermore, there is no requirement as to the *nature* of such a change (e.g., that the CR resemble the UR).

Backward conditioning as defined here does not specify the nature of the resulting learning. Issues such as whether the subject learns a backward association between the CS and the US or a forward association between the CS and some memory of the US are theoretical rather than empirical questions. As such, they are not relevant to an evaluation of the empirical evidence for backward conditioning.

Recent Demonstrations of Backward Conditioning

Although many of the earlier backward conditioning experiments did not include the necessary control procedures to provide strong evidence for the existence of backward conditioning, four recent experiments, including that of Keith-Lucas and Guttman (1975), appear to meet all of the necessary empirical criteria.

Keith-Lucas and Guttman

The experiment by Keith-Lucas and Guttman (1975) was a systematic replication of an earlier study by Hudson (1950). Hudson shocked rats while they were eating, then dropped a bundle of pipe cleaners into the chamber. In a subsequent test, the rats avoided the pipe cleaners. Keith-Lucas and Guttman repeated Hudson's study, using a toy hedgehog as the backward CS and including controls for nonpairing factors. They exposed four groups of rats to the following sequence of events. A novel stimulus panel containing a sugar pellet was inserted into one end of the chamber. As soon as the rat removed the pellet, a 750-msec, 1-mA foot-shock was administered, the chamber lights were extinguished, and the panel was removed. After a blackout of 1, 5, 10, or 40 sec (1-, 5-, 10-, and 40-sec backward groups), the lights were turned on and a toy hedgehog was presented for 1 min. Rats in two control groups underwent the same procedure except that exposure to either the shock (hedgehog group) or the hedgehog (shock group) was omitted.

All the subjects were tested 20 hours later. During the 10-min test, the panel and the hedgehog were presented at opposite ends of the chamber; their locations were reversed halfway through the test.

Five classes of behavior (location, normal approach, elongated approach, retraction, and pellet consumption) were recorded to assess avoidance of the panel (forward CS), the hedgehog (backward CS), or the shock location. This assessment revealed that the 1-, 5-, and 10-sec backward groups avoided the backward CS (hedgehog) rather than the forward CS (panel) or the shock location.

The 40-sec backward group and shock control group avoided the forward CS (striped panel) rather than the hedgehog; the hedgehog control group did not differentially avoid either stimulus.

Thus, following a single backward pairing with a US-CS interval of 10 sec or less, avoidance was conditioned to the CS (hedgehog) that followed shock. Factors other than stimulus pairing were clearly ruled out by the absence of comparable effects in the control groups and in the group that received a backward pairing with a 40-sec US-CS interval.

Wagner and Terry

Wagner and Terry (1975) hypothesized that manipulations that ensure the continued rehearsal of a US should result in more robust backward conditioning. Assuming that more rehearsal would occur with a "surprising" US than with an "expected" US, they designed an experiment to demonstrate that robust and enduring backward conditioning could be produced by maintaining the surprise value of the US over the course of the backward pairings. The surprise value of a US was established during preliminary discrimination training sessions in which one forward stimulus (CS⁺) was always followed by the US and another forward stimulus (CS⁻) was never followed by the US. Consequently, a US that followed the CS⁺ should be expected; a US that followed the CS⁻ should be unexpected or surprising. Backward pairings were then conducted with two new stimuli: one paired with an expected US, the other with a surprising US.

Wagner and Terry tested their hypothesis on rabbits in an eyelid conditioning paradigm. During initial discrimination sessions, a CS⁺ and CS⁻ consistently preceded the occurrence or nonoccurrence, respectively, of a 100-msec, 5-mA eye shock. Once the blink CRs consistently occurred to the CS⁺ but not to the CS⁻, 48 backward conditioning and testing sessions were conducted. For the experimental group, backward pairing trials and test trials, one of each per session, were embedded in the regular training trials. On some days, one backward CS (CS_E) was presented 500 msec after a CS⁺-US se-

quence, that is, after an expected US; on the remaining days, the other backward CS (CS_S) was presented 500 msec after a CS^- -US sequence, that is, after a surprising US. A 500-msec flashing light or vibrotactile stimulus served as the backward CSs; their designation as CS_E and CS_S was counterbalanced across subjects. A control group was included to assess the effect of simple exposure to a surprising US: Subjects received the same schedule of expected and surprising US trials, but the surprising US trials were not followed by a backward CS. For both groups, each session ended with a test trial in which either the CS_S or the CS_E was presented alone and the occurrence of blink CRs was recorded. The same criterion for recording blink CRs was used for both the forward and the backward CSs.

Wagner and Terry found clear evidence for backward conditioning to the CS that followed a surprising US. The subjects displayed significantly more blink CRs to the CS_S than to the CS_E —significantly more than the control subjects displayed in response to either stimulus. Moreover, these differences increased over the 24 tests with each stimulus.

Possible pseudoconditioning or sensitization effects of exposure to a surprising US were ruled out by the control subjects that received the same schedule of surprising USs but responded much less to the CS_S than did the rabbits in the experimental group. Furthermore, the counterbalanced pairing of one backward CS with a surprising US and another CS with an expected US can be viewed as a variation of the discriminative conditioning control procedure. The fact that the rabbits made significantly more responses to the CS_S than to the CS_E established that neither exposure to the backward CS prior to testing nor any forward conditioning effects with the US of the next trial could account for the results.

Heth and Rescorla

Heth and Rescorla (1973) investigated simultaneous and backward fear conditioning in rats using a conditioned punishment paradigm. Four groups of rats were exposed on consecutive days to two conditioning ses-

sions in which a total of 20 2-sec tone-light compound CSs and 20 4-sec, .5-MA shock USs were presented. Each day the rats in one group received 10 simultaneous pairings in which CS onset occurred 2 sec after US onset. The subjects in two other groups received 10 backward pairings each day; CS onset occurred 0 sec or .5 sec after US offset. The intertrial interval varied between 4 and 8 min. The control group received 20 CS presentations the 1st day and 20 US presentations the 2nd day.

During a subsequent test session, the punishing effects of the CS on previously established food-reinforced bar pressing were assessed. The degree of suppression produced by the response-contingent CS presentations served as the measure of conditioned fear. In an earlier experiment reported in the same article, 20 forward pairings of the same US and CS had resulted in a high degree of response suppression by the CS.

Although response suppression was greatest in the simultaneous condition, the bar pressing of the rats in the two backward groups was suppressed significantly more than that of the control rats during the first 10 min of the test. The two backward groups did not differ significantly from each other.

Mahoney and Ayres

Mahoney and Ayres (1976) investigated the fear-eliciting properties of a tone CS that had been paired a single time in a forward, simultaneous, or backward manner with a shock US. The suppression of licking by the thirsty rats during the presentation of the tone was the index of conditioned fear.

Baseline drinking rates were assessed for all six groups of rats before the conditioning session. During the conditioning session, the subjects in each group were presented with a single 4-sec tone CS and a 4-sec, 1-mA footshock US. For the two forward groups, tone offset preceded shock onset by 0 sec or 4 sec; for the simultaneous group, the tone and shock occurred simultaneously; for the two backward groups, tone onset followed shock offset by 0 or 4 sec. Rats in the sixth group served as control subjects; the tone was presented 3 min before shock onset. The effect of each of these pairings on licking

was assessed during 2 subsequent test days. The CS was presented after the 100th lick, and the time taken to complete 10 more licks was the dependent measure.

Although forward and simultaneous pairings produced more suppression than the backward pairing, the licking of the rats that had received a backward pairing was suppressed significantly more than was the licking of the control subjects. Subjects in the two backward groups displayed no significant difference in suppression.

Burkhardt (1980) replicated the major features of these results and demonstrated that the magnitude of suppression produced by a backward pairing increased as US intensity increased.

Summary

Each of the preceding experiments satisfies all of the empirical criteria and thus qualifies as a demonstration of backward conditioning. In each experiment a backward US-CS pairing produced a statistically significant change in behavior that resembled the effect of forward pairing of the same stimuli. Each studies included control procedures that ruled out the possibility that the change in behavior was due to factors other than the backward pairings, such as pseudoconditioning or sensitization. Finally, in all four studies the possibility that the effect was due to a delayed forward pairing of the CS with the US of a following trial was ruled out: Heth and Rescorla (1973) used long intertrial intervals (4–8 min); Wagner and Terry (1975) used both long intertrial intervals and a discriminative conditioning control procedure; and in two studies (Keith-Lucas & Guttman, 1975; Mahoney & Ayres, 1976), backward conditioning occurred after a single pairing.

Early Studies of Backward Conditioning

The discrepancy between the recent demonstrations of backward conditioning and the prevalent view that backward conditioning does not occur suggests that a reevaluation of the early studies of backward conditioning, which serve as a basis for the current view, is warranted.

Pavlov (1927/1960, p. 27) initially concluded that CRs could not be established if the CS was presented after the US. This conclusion was based on unpublished experiments by Krestovnikov in which dogs were given many (374 or 427) US-CS pairings; in a subsequent test, the CS presented alone failed to elicit a CR. In a later lecture (pp. 391–394), however, Pavlov stated that his earlier position was a “probable error” (p. 391). He described the results of experiments by Pavlov, Kreps, Podkopaev, Prokrov, and Koupalov in which only a few backward pairings were administered: “The hitherto neutral stimulus when now tested alone revealed undoubted conditioned properties” (p. 393). He then suggested that since Krestovnikov’s experiments assessed only the effects of multiple trials, his earlier conclusions represented

a brilliant illustration of the danger of too hasty generalizations. We imagined that if it were a true conditioned reflex it would increase in intensity upon repetition of the combination, and not diminish and finally vanish as happens in these experiments. (p. 394)

Still later, Pavlov (1928, p. 381) restated his conclusion that weak CRs develop as a result of initial backward pairings, adding that with repeated pairings the CS becomes inhibitory. Thus Pavlov’s final position was not that the backward procedure produces no conditioning but rather that it has “a double effect: at first temporarily it assists in the formation of the conditioned reflex, and then destroys it” (1932, p. 93).

The existence of the “double effect” first reported by Pavlov has been confirmed in subsequent studies. Several investigators have reported the appearance of CRs following a few backward pairings (e.g., Switzer, 1930; Wolffe, 1930), whereas those assessing the effects of multiple backward pairings have found either no conditioning (e.g., Bernstein, 1934; Porter, 1938) or conditioned inhibition (e.g., Moscovitch & Lo-Lordo, 1968). Spooner and Kellogg (1947) administered different numbers of backward pairings and reported that CRs appeared following initial pairings but diminished with repeated pairings. Heth (1976) also reported that backward pairings produced a biphasic acquisition function, with the CS

controlling excitatory tendencies after 10 pairings and inhibitory tendencies after 160 pairings. Unfortunately, many of the early studies did not include the necessary controls to rule out the possibility that factors other than stimulus pairing were producing the apparent CRs during the initial backward pairings. A few studies, however, did include such controls.

Mowrer and Aiken (1954), for example, demonstrated backward conditioning using a conditioned punishment paradigm. One group of rats received a single backward pairing of a shock and light on each of 5 successive days. On the 6th day the light was made contingent on a previously established bar-press response; suppression of this response served as a measure of conditioning. Suppression in these rats was significantly greater than rats that had received unpaired CS and US presentations. Matsumiya (1960) replicated these results. The Heth and Rescorla (1973) study was in fact a systematic replication of these earlier studies.

Two studies have demonstrated backward conditioning of the galvanic skin response (GSR). In both, backward pairing produced significantly greater changes in the GSR than did unpaired presentations of the CS and US (Champion & Jones, 1961; Zeiner & Grings, 1968). Thus empirical evidence for backward conditioning does exist in the early literature, and this evidence is not only consistent with but anticipates the findings of recent backward conditioning studies.

The frequent reports in the early literature of unsuccessful backward conditioning experiments also deserve attention. As mentioned earlier, several experimenters failed to observe CRs following multiple backward pairings (e.g., Pavlov, 1927/1960; Porter, 1938). Others who reported that apparent CRs occurred early in training concluded that these were not genuine CRs because their occurrence declined with repeated pairings (e.g., Spooner & Kellogg, 1947). And some experimenters (Grether, 1938; Prokasy, Hall, & Fawcett, 1962) reported that backward groups do not differ from pseudoconditioning groups. Because the results of these studies are frequently cited to support the conclusion that backward conditioning is not a genuine phenomenon (Cau-

tela, 1965; Gormezano & Moore, 1969; Mackintosh, 1974), it is worthwhile to carefully examine the evidence provided by some of them.

An experiment by Spooner and Kellogg (1947) is probably the most commonly cited. Spooner and Kellogg compared forward, simultaneous, and backward conditioning of the finger withdrawal response in human subjects. Three features of their data led to their conclusion that backward and simultaneous conditioning are entirely different phenomena from forward conditioning.

1. *Differences in response latency.* Spooner and Kellogg found that the CR latencies of the backward and simultaneous groups were shorter and more variable than those of the forward groups and suggested that these differences reflect "a basic distinction between the two processes" (Spooner & Kellogg, 1947, p. 331). When one examines the mean latencies for all groups, however, one finds even greater differences in latencies between forward groups with different interstimulus intervals. As Pavlov (1927/1960, p. 88) reported, response latency varied as a direct function of the interval between the CS and the US. Thus shorter latencies in the backward and simultaneous groups may simply represent an extension of this general principle. Nevertheless, if differences in response latency constitute evidence for different processes, one is forced to conclude that different processes also underlie forward conditioning at different interstimulus intervals.

2. *Differences in shock threshold.* The US intensity was adjusted for individual subjects throughout the experiment to prevent adaptation to the shock, and it was found that the US intensity required to elicit unconditioned finger withdrawal was higher for the backward and simultaneous groups than for the forward groups. Although Spooner and Kellogg (1947) presented this as evidence that forward and backward conditioning are different phenomena, no statistical analysis was performed to test the significance of these differences, and no evidence was offered that the US threshold is related in any way to the conditioning process. Moreover, because these differences were present from the first test sessions, it is likely

that the groups were not evenly matched with respect to initial threshold levels.

3. *Differences in progress.* When the frequency of CRs was plotted as a function of trials, differences between the forward, backward, and simultaneous groups emerged. Whereas the forward groups showed an increase in CR frequency as a function of training, the backward and simultaneous groups showed a decrease and eventual cessation of the CR with repeated pairings. Essentially, these results confirm those reported by Pavlov. Unlike Pavlov (1927/1960, p. 393), however, Spooner and Kellogg (1947) concluded that backward and simultaneous conditioning were not "true conditioning." This conclusion was reached in spite of the fact that their experiment included no control groups to assess the effects of factors other than stimulus pairing. Without such controls, it is impossible to determine whether the CRs produced by the backward pairings reflect genuine conditioning. Nevertheless, their tenuous conclusion that backward conditioning is a fundamentally different process from forward conditioning frequently has been cited (e.g., Osgood, 1953; Terrace, 1973).

Subsequent authors further complicated the issue by exaggerating the claims of Spooner and Kellogg (1947) regarding the possibility that backward conditioned CRs might be due to pseudoconditioning. Spooner and Kellogg's actual conclusion was that "it is possible that backward and simultaneous CRs are special cases of pseudoconditioning, although our data do not wholly support such a view" (p. 334). Cautela (1965) claimed, however, that "they concluded that backward-conditioned responses probably are special cases of pseudoconditioning" (p. 139), and Gormezano and Moore (1969) stated that Spooner and Kellogg "clearly articulated the notion that forward and backward conditioning are fundamentally different processes with the latter being a form of pseudoconditioning" (p. 145). Surprisingly, Bugelski (1956) claimed that Spooner and Kellogg found that "any responses that appear early in training might be attributed to sensitization" (p. 129), even though there was no mention of sensitization in Spooner and Kellogg's article.

An earlier experiment by Grether (1938) has also been referenced frequently to support the view that backward conditioning may be attributed to pseudoconditioning. Grether presented two monkeys with 10 backward pairings of a US and a CS. For one subject, the US was flashlight powder that exploded 2 feet (.6 m) in front of the monkey; for the other, the US was a snake "blowout" that uncurled to 6 inches (15.24 cm) from the monkey's nose and produced a rattling noise. For both, the sound of an electric bell served as the CS. Two additional monkeys served as pseudoconditioning control subjects and received 10 presentations of the US alone (exploding powder for one; the snake blowout for the other). When subsequently tested with the bell, Grether reported, all subjects "reacted violently."

Although, as Grether (1938) suggested, this experiment demonstrates the need for pseudoconditioning control groups, it by no means provides evidence against the possibility that backward conditioning is genuine conditioning. First, "reacting violently" is not a quantifiable measure with which to make fine distinctions between the groups. Second, because the control subjects apparently responded maximally in terms of this qualitative measure, the comparison between the backward and pseudoconditioning treatments may have suffered from a ceiling effect. The possibility of such a ceiling effect cannot be ruled out in the absence of evidence that any experimental treatment, such as forward pairings, would produce greater responding than that shown by the pseudoconditioning subjects.

In fact, Prokasy et al. (1962) included such a forward conditioning group in a subsequent failure to demonstrate differences between backward conditioning and pseudoconditioning groups. They compared GSR conditioning for forward, simultaneous, backward, randomly paired, US-only, and CS-only groups. None of the conditioning groups, including the forward one, differed significantly from the groups that had received random pairings or US-alone presentations. Nevertheless, even these results have been used to argue that backward conditioning does not exist (e.g., Mackintosh, 1974).

In summary, these experiments by Grether (1938), Spooner and Kellogg (1947), and Prokasy et al. (1962), which have been cited frequently as evidence that backward conditioning does not exist, in fact provide little if any support for the view that backward conditioning is not possible.

Previous Reviews of Backward Conditioning

Our evaluation of the earlier empirical evidence for backward conditioning suggests a conclusion that is similar to that reached by Razran (1956) in his review of the Russian and American literature but is clearly at odds with the influential conclusions of Cautela (1965) and Mackintosh (1974). This discrepancy appears to stem in part from differences in the criteria used.

Cautela

Cautela (1965) evaluated the evidence for backward conditioning in terms of two theoretical views of conditioning: a "substitute stimulus" definition and a "modifying stimulus" definition. He used criteria derived from the former definition to evaluate classical conditioning studies and criteria derived from the second to evaluate studies that used operant conditioning procedures to assess the effect of stimulus pairing.

According to the substitute stimulus view, conditioning is said to have taken place only if a neutral stimulus, paired with a stimulus that always elicits a particular response, elicits that same response: The response to the CS must resemble the response to the US. Cautela required as well that CR probability "increase with repeated reinforcements and remain relatively stable when some maximum strength is reached. . . . Unstable or temporary conditioned responses would be attributed to pseudoconditioning" (p. 136).

Because backward conditioning is maximal following the initial trials and may disappear with repeated pairings (Spooner & Kellogg, 1947), it is readily apparent why Cautela's view of conditioning would lead to the conclusion that backward conditioning does not occur. Indeed, after reviewing the

classical conditioning studies, Cautela concluded that

only one study seems unquestionably to have fulfilled the criteria for the substitute stimulus definition of conditioning. In all other studies, the conditioned responses were unstable, did not become stronger with practice or were called pseudoconditioning by some investigator. (p. 141)

Cautela then reviewed studies that used an operant conditioning procedure to assess the effect of prior backward pairings of a CS and a US, basing his evaluation on criteria determined by the modifying stimulus definition of conditioning. According to this view, conditioning is "the process of modifying the effects of a stimulus on the behavior of organisms by associating it with a second stimulus. . . . The stimulus can inhibit as well as facilitate response probability" (p. 136). This definition is less restrictive than our definition of backward conditioning because it encompasses demonstrations of inhibitory conditioning (i.e., behavioral change produced by backward pairings opposite to that produced by forward pairings of the CS and US). Five of the six such operant studies reviewed by Cautela provided evidence of backward conditioning, according to his modifying stimulus criteria. Of these, four were demonstrations of inhibitory conditioning.

In the final section of Cautela's review, it becomes clear that, in addition to the criteria made explicit from the start, Cautela had some implicit criteria based on assumptions about the nature of the association resulting from backward pairings. Six studies met Cautela's explicitly stated criteria for backward conditioning (the one substitute-stimulus study and five modifying-stimulus studies). Yet, rather than accepting the evidence from them, Cautela dismissed all six on the grounds that all had used noxious stimuli as the US. He first suggested that "the apparent necessary involvement of a noxious stimulus in backward conditioning supports the sensitization hypothesis" (p. 141). He then proposed an alternative "duration-of-pain" hypothesis to account for the appearance of CRs following backward US-CS pairings: "Cases of reported backward conditioning with the use of a noxious stimulus are really

cases of simultaneous conditioning of the CS and pain perception" (p. 141).

Interestingly, Cautela proposed these hypotheses to explain the appearance of CRs following a backward pairing, even though four out of six of the studies that met his criteria for backward conditioning actually demonstrated conditioned *inhibition*. Thus, although his duration-of-pain hypothesis might explain at a theoretical level the development of backward conditioning, it does not explain why repeated backward pairings would produce an effect that is opposite to that produced by repeated forward pairings.

Cautela concluded, "Before either the sensitization hypothesis or the duration-of-pain hypothesis can be ruled out, backward conditioning will have to be obtained without the use of a noxious stimulus" (p. 143). This conclusion is erroneous on two accounts. First, even when noxious stimuli are used, sensitization can be ruled out by the use of sensitization control groups. Second, the duration-of-pain hypothesis is a theoretical explanation of the results of backward pairings. As we state in the introduction, an empirical demonstration of backward conditioning does not require evidence that the association learned be backwards. If one were to make such a requirement, backward conditioning could not be demonstrated with certainty, regardless of the stimulus used. Any stimulus, noxious or rewarding, can be hypothesized to produce perceptual after-effects, the duration of which for an animal cannot be determined with certainty. Thus even a demonstration of backward conditioning with an appetitive US would not rule out the possibility that the underlying association was a forward one between the CS and aftereffects of the US.

Although Cautela's review appeared prior to many of the more well-controlled demonstrations of backward conditioning, it is apparent that none of the experimental evidence for backward conditioning to date would satisfy both his explicit and his implicit criteria. In fact, not even widely accepted demonstrations of forward conditioning (e.g., Peterson, Ackil, Frommer, & Hearst, 1972; Wasserman, 1973) meet all of the criteria derived from the stimulus-substitution definition of conditioning.

Mackintosh

Mackintosh (1974) did not make explicit the criteria he used. He gave two reasons, however, for rejecting the apparently successful backward conditioning studies: (a) the numerous failures to replicate successful experiments and (b) the problems arising from the use of the GSR as a CR.

In view of the number of successful demonstrations of backward conditioning, failures to replicate successful studies do not constitute sufficient evidence for questioning its existence, especially when several such failures (e.g., Grether, 1937; Prokasy et al., 1962) were poorly controlled or provided little direct evidence against the effectiveness of backward conditioning or both. The problems with the GSR, however, do deserve some consideration.

Mackintosh stated that, in a test for backward conditioning,

the CS presented alone is a novel event, and the occurrence of a GSR on the first few test trials is most simply thought of as a component of the orientation reaction. In support of this . . . conditioning is maximal on the first few test trials and thereafter tends to decline. (p. 59)

Although this argument is potentially viable, a close examination of the studies that Mackintosh cited to support his interpretation reveals that this possibility may have been ruled out by the control data reported in the studies. For example, Champion and Jones (1961) reported conditioning of the GSR following a small number of backward or forward pairings of a shock and a tone. Three groups of people received seven forward CS-US pairings, seven backward US-CS pairings, or seven unpaired presentations of the US and CS. The forward and backward groups were tested with the CS alone following one, three, five, and seven pairings. The subjects in both groups responded more on all tests than did the unpaired control subjects, who were tested following the same number of US and CS presentations. Because the contribution of orientation reactions to the GSR should have been the same in all groups, the higher incidence of responding in the backward group compared with the unpaired group cannot be attributed to orientation reactions alone.

In a footnote Mackintosh stated:

An additional complication (deriving from the use of students as subjects) has been noted by Zeiner and Grings (1968): subjects that reported expecting the CS to signify something showed a markedly higher level of GSR responding than did less suspicious subjects. (p. 59)

Zeiner and Grings first replicated the Champion and Jones (1961) study and found significantly greater GSR CRs in subjects who had received backward pairings than in subjects in an unpaired control group. They then divided the backward group into two subgroups on the basis of verbal reports and responses to a questionnaire regarding the perceived significance of the CS. A post hoc comparison of the data of these subgroups revealed that subjects who attributed significance to the CS had displayed greater GSR CRs than had those who did not think the CS was significant. This is an interesting correlation, but it has no bearing on the original finding that subjects in the backward group responded more than those in the unpaired control group. First, because the assessment of perceived significance of the CS was conducted after the conditioning process, it is impossible to determine whether subjects who attached more significance to the CS were more suspicious to start with or whether they perceived the CS as significant because of the conditioning process. Even if one wished to infer from this finding that suspicious subjects are more easily conditioned, this would not constitute evidence against the efficacy of backward pairings in producing such conditioning.

Razran

In his review of Soviet and American backward conditioning experiments, Razran (1956) concluded:

On the whole, the analyzed evidence is unmistakable in demonstrating that B.C. [backward conditioning] is not a case of pseudoconditioning, but is a genuine CR-associative manifestation, and that stable backward CR's can be obtained and maintained under favorable experimental conditions. (p. 67)

Razran did not enumerate the criteria he used to evaluate the evidence that led to this conclusion. In fact, it appears that his only criteria were that a backward procedure was

used and that the backward procedure resulted in the appearance of CRs. He accepted as evidence of genuine backward conditioning all studies that observed CRs as a result of backward pairings, regardless of whether controls for factors other than the backward pairings were included. In fact, several of the studies that Razran cited as the most conclusive demonstrations of backward conditioning (Spooner & Kellogg, 1947; Wolfle, 1930, 1932) did not include the necessary control groups to rule out pseudoconditioning or sensitization effects. Thus, although we agree with Razran's conclusion that backward conditioning is a genuine phenomenon, we do not agree with the evidence on which his conclusion was based. Razran's unqualified acceptance of evidence from uncontrolled studies is probably why his review has had little impact on views of backward conditioning in the American literature.

Possible Frameworks for Backward Conditioning

The existence of backward conditioning is incompatible with several traditional views of classical conditioning (e.g., Hull, 1943) and poses problems for contemporary views that emphasize the predictive function of the CS in the process of learning—for example, the information hypothesis (Egger & Miller, 1962) or the contingency theory (Rescorla, 1967). Thus the empirical evidence for backward conditioning provides a theoretical challenge that should stimulate new or modified views of conditioning processes.

On the other hand, a few contemporary perspectives do seem able to account for backward conditioning. For example, the biphasic effect of backward pairings, with excitatory conditioning occurring after initial pairings and inhibition occurring after repeated pairings, is consistent with Wagner and Terry's (1975) postepisodic rehearsal theory. They proposed that associations between the CS and US depend on the simultaneous rehearsal of the two events in short-term memory and that more post-US rehearsal occurs when the US is surprising. During initial backward pairings, the US is relatively surprising, and post-US rehearsal would be in progress when the CS is pre-

sented; excitatory conditioning would occur because of the joint rehearsal of the two events. On later trials the US would become expected on the basis of contextual cues, and post-US rehearsal would be attenuated. Then CS rehearsal in the absence of US rehearsal would cause the CS to undergo extinction, and with repeated pairings the CS might acquire inhibitory properties.

Both the existence of backward conditioning and some of its prominent features might also be viewed profitably in terms of recent approaches that emphasize the role for phylogenetic factors in learning. Fantino and Logan (1979) summarized the major premises that underly this "natural history" perspective:

(1) Intricate interactions occur between ontogenetic and phylogenetic factors in learning. The nature of these interactions is determined by the ecological demands placed on the species in question. (2) Learning processes occur with tremendous diversity, which is governed by the nature of the situation that renders learning of adaptive benefit. (3) Only by understanding how learning functions to maximize reproductive success in the organism's natural environment can we arrive at complete explanations of the nature of behavior change. For it is here that the demands for survival are defined. (p. 386)

Most successful demonstrations of backward conditioning to date have several common features: the use of noxious stimuli as USs, the small number of US-CS pairings, the unpredictability of the US or all three. Keith-Lucas and Guttman (1975) demonstrated particularly robust backward conditioning in rats using a CS similar to a natural predator. The contribution of such features to the effectiveness of backward pairings would be expected if the process that mediates backward conditioning evolved as an adaptation to attacks by conspecifics or predators. An animal that could associate sudden, aversive stimulation with a subsequent novel stimulus would gain clear reproductive advantage.

Regardless of the particular theoretical framework in which backward conditioning is considered, the empirical evidence for its existence can no longer be ignored. Thus the time for disputing whether backward conditioning is possible is past; it is time, instead, for systematic exploration of the conditions under which it occurs and the

variables that affect the magnitude and duration of the effect.

References

- Bernstein, A. L. Temporal factors in the formation of conditioned eyelid reactions in human subjects. *Journal of General Psychology*, 1934, 10, 173-197.
- Bugelski, B. R. *The psychology of learning*. New York: Holt, 1956.
- Burkhardt, P. E. One-trial backward fear conditioning in rats as a function of US intensity. *Bulletin of the Psychonomic Society*, 1980, 15, 9-11.
- Cautela, J. R. The problem of backward conditioning. *Journal of Psychology*, 1965, 60, 135-144.
- Champion, R. A., & Jones, J. E. Forward, backward, and pseudoconditioning of the GSR. *Journal of Experimental Psychology*, 1961, 62, 58-61.
- Egger, M. D., & Miller, N. E. Secondary reinforcement in rats as a function of information value and reliability of the stimulus. *Journal of Experimental Psychology*, 1962, 64, 97-104.
- Fantino, E., & Logan, C. A. *The experimental analysis of behavior: A biological approach*. San Francisco: Freeman, 1979.
- Gormezano, I., & Moore, J. W. Classical conditioning. In M. H. Marx (Ed.), *Learning: Processes*. Toronto, Ontario, Canada: Collier-Macmillan, 1969.
- Grant, D. A., & Adams, J. K. "Alpha" conditioning in the eyelid. *Journal of Experimental Psychology*, 1944, 34, 136-147.
- Grether, W. F. Pseudoconditioning without paired stimulation encountered in attempted backward conditioning. *Journal of Comparative Psychology*, 1938, 25, 91-96.
- Groves, P. M., & Thompson, R. F. Habituation: A dual-process theory. *Psychological Review*, 1970, 77, 419-450.
- Heth, C. D. Simultaneous and backward fear conditioning as a function of a number of CS-UCS pairings. *Journal of Experimental Psychology: Animal Behavior Processes*, 1976, 2, 117-129.
- Heth, D. C., & Rescorla, R. A. Simultaneous and backward fear conditioning in the rat. *Journal of Comparative and Physiological Psychology*, 1973, 82, 434-443.
- Hudson, B. B. One-trial learning in the domestic rat. *Genetic Psychology Monographs*, 1950, 41, 99-145.
- Hull, C. L. *Principles of behavior*. New York: Appleton-Century-Crofts, 1943.
- Hulse, S. H., Deese, J., & Egerth, H. *The psychology of learning*. New York: McGraw-Hill, 1975.
- Kalish, H. I. Strength of fear as a function of the number of acquisition and extinction trials. *Journal of Experimental Psychology*, 1954, 47, 1-9.
- Keith-Lucas, T., & Guttman, N. Robust single-trial delayed backward conditioning. *Journal of Comparative and Physiological Psychology*, 1975, 88, 468-476.
- Kimble, G. A. *Hilgard and Marquis' conditioning and learning*. New York: Appleton-Century-Crofts, 1961.
- Logue, A. W. Taste aversion and the generality of the laws of learning. *Psychological Bulletin*, 1979, 86, 276-296.

- Mackintosh, N. J. *The psychology of animal learning*. New York: Academic Press, 1974.
- Mahoney, W. J., & Ayres, J. J. B. One-trial simultaneous and backward fear conditioning as reflected in conditioned suppression of licking in rats. *Animal Learning and Behavior*, 1976, 4, 357-362.
- Matsumiya, Y. The effects of US intensity and CS-US pattern on conditioned emotional response. *Japanese Psychological Research*, 1960, 2, 35-42.
- Millenson, J. R., & Leslie, J. C. *Principles of behavioral analysis*. New York: Macmillan, 1979.
- Moeller, G. The CS-UCS interval in GSR conditioning. *Journal of Experimental Psychology*, 1954, 48, 162-166.
- Moscovitch, A., & LoLordo, V. M. Role of safety in the Pavlovian backward fear conditioning procedure. *Journal of Comparative and Physiological Psychology*, 1968, 66, 673-678.
- Mowrer, O. H., & Aiken, F. G. Contiguity vs. drive-reduction in conditioned fear: Temporal variations in conditioned and unconditioned stimulus. *American Journal of Psychology*, 1954, 67, 26-38.
- Osgood, C. E. *Method and theory in experimental psychology*. New York: Oxford University Press, 1953.
- Pavlov, I. P. *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (G. V. Anrep, Ed. and trans.). New York: Dover, 1960. (Originally published, 1927.)
- Pavlov, I. P. *Lectures on conditioned reflexes* (W. H. Gantt, trans.). New York: International, 1928.
- Pavlov, I. P. The reply of a physiologist to psychologists. *Psychological Review*, 1932, 39, 91-127.
- Peterson, G. B., Ackil, J. E., Frommer, G. P., & Hearst, E. Conditioned approach and contact behavior towards signals for food or brain-stimulation reinforcement. *Science*, 1972, 177, 1009-1011.
- Porter, J. M., Jr. Backward conditioning of the eyelid response. *Journal of Experimental Psychology*, 1938, 23, 403-410.
- Prokasy, W. F., Hall, J. F., & Fawcett, J. T. Adaptation, sensitization, forward and backward conditioning, and pseudoconditioning of the GSR. *Psychological Reports*, 1962, 10, 103-106.
- Razran, G. Backward conditioning. *Psychological Bulletin*, 1956, 53, 55-69.
- Rescorla, R. A. Pavlovian conditioning and its proper control procedures. *Psychological Review*, 1967, 74, 71-80.
- Rozin, P., & Kalat, J. W. Specific hungers and poison avoidance as adaptive specializations of learning. *Psychological Review*, 1971, 78, 459-486.
- Schwartz, B. *Psychology of learning and behavior*. New York: Norton, 1978.
- Spence, K. W., & Runquist, W. N. Temporal effects of conditioned fear on the eyelid reflex. *Journal of Experimental Psychology*, 1958, 55, 613-616.
- Spooner, A., & Kellogg, W. N. The backward conditioning curve. *American Journal of Psychology*, 1947, 60, 321-334.
- Switzer, C. A. Backward conditioning of the lid reflex. *Journal of Experimental Psychology*, 1930, 13, 76-97.
- Terrace, H. S. Classical conditioning. In J. A. Nevin & G. S. Reynolds (Eds.), *The study of behavior: Learning, motivation, emotion, and instinct*. Glenview, Ill.: Scott, Foresman, 1973.
- Wagner, A. R., & Terry, W. S. Backward conditioning to a CS following an expected vs. a surprising UCS. *Animal Learning and Behavior*, 1975, 3, 370-374.
- Wasserman, E. A. Pavlovian conditioning with heat reinforcement produces stimulus directed pecks in chicks. *Science*, 1973, 181, 875-877.
- Wolfe, H. M. Time factors in conditioning finger withdrawal. *Journal of General Psychology*, 1930, 4, 372-378.
- Wolfe, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *Journal of General Psychology*, 1932, 7, 80-103.
- Zeiner, A., & Grings, W. W. Backward conditioning: A replication with emphasis on conceptualizations by the subject. *Journal of Experimental Psychology*, 1968, 76, 232-235.

Received January 30, 1980 ■