Spontaneous Recovery

Robert A. Rescorla
Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Spontaneous recovery from extinction is one of the most basic phenomena of Pavlovian conditioning. Although it can be studied by using a variety of designs, some procedures are better than others for identifying the involvement of underlying learning processes. A wide range of different learning mechanisms has been suggested as being engaged by extinction, most of which have implications for the nature of spontaneous recovery. However, despite the centrality of the notion of spontaneous recovery to the understanding of extinction, the empirical literature on its determinants is relatively sparse and quite mixed. Its very ubiquity suggests that spontaneous recovery has multiple sources.

Experimental extinction is one of the fundamental observations of Pavlovian conditioning. Just as the arranging of a positive relation between a conditioned stimulus (CS) and an unconditioned stimulus (US) produces acquisition of conditioned responding, breaking that relation produces extinction of that responding. However, similar to many terms in the behavioral sciences, the word “extinction” is used in at least three different senses: as a procedure, as a result, and as an explanation. If we are to understand extinction experiments, it is extremely important that we keep these senses distinct from each other.

One use of the term is as an experimental procedure or independent variable under the control of the experimenter, as when one says, “Following learning, we subjected the animal to an extinction procedure.” Most frequently, this is meant to refer to a procedure in which the original conditions of learning are disrupted. The most common extinction procedure consists of presenting a stimulus alone, so that it now fails to signal the outcome. However, other procedures, such as retaining the US but arranging for it to be independent of the CS are also available and of interest (see Rescorla 2001a).

Another use of the term is as an experimental result or dependent variable under the control of the animal, as when one says, “When the stimulus was presented alone, the behavior extinguished.” The prototypical example is one in which responding that was established by training deteriorates, often to a level such as that prior to learning.

A third use of the term extinction is as a process or intervening variable that is intended to provide an explanation, as when one says, “When we arranged for the stimulus to be presented alone, the behavior deteriorated because of extinction.” Normally, it is this process that is of interest. We would like to understand the basis of the change in behavior resulting from the change in procedure, whether that understanding is achieved at a conceptual or a neural level. Consequently, throughout this article the term extinction will refer to the learning process inferred when the procedure produces a particular result. When there is the possibility of misunderstanding, the phrase “extinction process” will be used.

Because interest primarily centers on the learning process that occurs as a result of an extinction procedure, it is important to separate that learning from a wide variety of other effects that govern performance. The issue here is analogous to that of understanding the learning that occurs during an acquisition procedure. Elsewhere we have argued that the measurement of learning demands attention to two points in time: t1, during which the opportunity to learn is given, and then a separate t2, during which an assessment is made of that learning. (Rescorla and Holland 1976; Rescorla 1988). The comparison that indicates that learning has occurred is that between two animals (or two stimuli or responses) given a common t2 test following different opportunities for learning at t1. This comparison is superior to the common alternative of examining responding during t1, at which the animals are receiving different learning treatments. Data taken during t1 necessarily confound the differences in the current circumstances under which learning is assessed with differences in the learning that prior treatments might have produced. We have argued that for this reason acquisition curves are in fact deeply flawed as a way to measure learning.

A similar point applies to the learning that occurs in extinction. In this case, we need to administer a common test for stimuli or animals given different extinction experiences, as indicated in the first portion of Figure 1. In the simplest case, we can compare responding to two stimuli (S1 and S2) at a common t2 after both have had the same initial acquisition but then differ in whether or not they were given extinction at t1. Differences in t2 test performance would then index the differences in learning that occurred at t1. Clearly, comparisons between stimuli during the t1 extinction experience are of limited value because any differences might be the product of the current conditions of testing rather than of the memory for the learning that has occurred. That is, extinction curves are of very limited use in understanding the underlying process.

In this context, the phenomenon of spontaneous recovery has a complex role. That phenomenon suggests that the results that we obtain in a t2 assessment may be quite different depending on the length of time that intervenes between the t1 extinction experience and the t2 test.

It has been known since Pavlov’s (1927) early experiments that the loss of behavior that results from presenting the stimulus alone at t1 is not entirely permanent. Rather, with the passage of time following nonreinforcement, there is some “spontaneous recovery” of the initially learned behavior. Introducing greater time delays between t1 extinction treatment and t2 test provides the opportunity for greater spontaneous recovery.

At minimum, the phenomenon of spontaneous recovery provides some information about what process fails to account for the loss of behavior when an extinction procedure is conducted. It suggests that the loss does not simply involve the removal of what was learned in acquisition. As Pavlov noted, if an extinction procedure had erased the acquisition learning, then there would be no basis for behavior to return with time. It suggests that instead something happens during the extinction pro-
ing at the end of an extinction session and at the beginning of a test is commonly inferred from a comparison between responding that was depressed during an extinction session and responding to one that has never been trained prior to extinction but is still given the same interval between extinction and test sessions in which to “recover.” Otherwise, it is possible that the increase in responding represents a general tendency to increase responding with time independently of the original acquisition; that is, it might not reflect recovery of the initial learning at all. In fact, there are only a few studies (see Robbins 1990) that have deliberately made such a comparison. But any detailed investigation of spontaneous recovery should be sensitive to this possibility.

To justify the second assumption (that the increase in responding represents dissipation of the extinction learning), comparison should be made with a stimulus that has trained, but not been extinguished, and that evokes responding at a level comparable to that of the extinguished stimulus before the delay interval. For instance, one might compare, for the same animal, changes with time in a trained and extinguished stimulus and a stimulus that is only partially trained, so as to attain the same response level. It is only if the former stimulus shows greater growth with time that one would conclude that there is a loss of the learning that occurred during extinction, as distinct from a general change in performance for any previously trained stimulus showing behavior at that level.

Aside from the need to occasionally include these comparisons, the spontaneous recovery design shown in Figure 1B has the drawback of repeated testing with the same stimulus, with the consequence that different numbers of extinction trials have necessarily preceded the trials being compared. A better design, which is also frequently used, is shown in Figure 1C. In that design two stimuli are both trained and extinguished but then given different amounts of time to recovery prior to the test. This design avoids repeated testing on the same stimulus and has the advantage of an explicit test session. But unfortunately, the tests of the two stimuli differ not only in the time since extinction but also in the time since original training and in the overall test context and age of the animal.

To avoid these confoundings, we have frequently adopted a somewhat different design for studying spontaneous recovery, as illustrated in Figure 1D. In this comparison, two groups of animals receive conditioning followed by extinction and a test. The groups differ in the placement of the extinction with regard to the test. For one group (S1) the extinction occurs at a temporal distance from the test, so as to allow spontaneous recovery. For the other group (S2), extinction occurs immediately prior to test, minimizing recovery. The evidence for recovery is then the difference in responding at the time of the common test. That allows comparison of responding during the same test session, after the same number of extinction trials, to stimuli that share the time since original training but differ in the time since their extinction.

The design can be further strengthened if the two stimuli are both trained in the same animal. An especially important advance is the ability to control for general changes in the animal’s state or to differential similarity of the test conditions to those of original training. One difficulty with the standard procedure for assessing spontaneous recovery is that one part of the comparison comes from a session in which the animal is undergoing an extinction procedure and is therefore likely to
be experiencing a variety of new stimulus events that may have various emotional consequences. The presence of those new events could, in themselves, artificially depress responding to a lower level than that which is warranted by the learning that is occurring. The likely absence of those stimuli at the beginning of the test session could allow greater responding for reasons that are not central to the learning that is occurring in extinction. But the final design avoids such differences in the context of performance for the stimuli being compared.

By using within-subject versions of this last procedure, Rescorla (1997a,b) has shown evidence of substantial spontaneous recovery in a variety of conditioning preparations. An example is shown in Figure 2 for a Pavlovian magazine approach situation with rats. In this preparation, S1 and S2 are counterbalanced as a 30-sec light and noise, each paired with the delivery of food and then given nonreinforcement. The result is an increase and decrease in anticipatory responding involving investigation of the food delivery site, the magazine. A test then takes place either immediately after the last extinction session or after a delay. The comparison of interest is that between S1 and S2 in the common test session. Figure 2 clearly indicates greater recovery for the stimulus extinguished first, S1. One may note that S1 also shows more responding in the test than it did at the end of extinction; but as noted above, this comparison is flawed.

Of course, no procedure is without its shortcomings. In this alternative procedure for assessing spontaneous recovery, one of the stimuli must necessarily be extinguished before the other. Consequently, the order in which the stimuli are extinguished, and the interval between training and test, are both confounded with the interval of interest, between extinction and test. Although application of the design has not revealed any resulting differences in the course of extinction itself (Fig. 2), one must still be cautious about the implications of such a confounding. Consequently, it may be most wise to use both procedures C and D in any research program investigating spontaneous recovery.

Bases for Spontaneous Recovery

Spontaneous recovery is such a widespread phenomenon, both in terms of the variety of learning paradigms in which it occurs and the frequency with which it is reported, that it would be surprising if it had only one source. Indeed, many different sources have been suggested, most of which have received some empirical support. But not all of these are of equal interest in helping us understand the learning involved in extinction. Here, I attempt a rough categorization of those sources.

Local Performance Effects

As noted above, when one makes the standard comparison of responding on the final trials from extinction with that on the initial trials in a test session, one can expect that they will differ in a number of ways unrelated to the learning that occurs in extinction. For instance, the repeated evocation of a conditioned response might lead to temporary fatigue that adds to the loss of behavior. It is not uncommon to see behavior decline over the course of a conditioning session even when reinforcement continues (see McSweeney and Swindell 1999). Similarly, the surprising deletion of anticipated reinforcement may well lead to emotional responses that have a disruptive effect on performance. For instance, the surprising omission of food can be expected to lead to frustration, which changes the stimulus environment and may evoke responses of its own (see Amsel 1958). Effects such as these might well be expected to dissipate before the beginning of the test session, hence revealing any incompleteness of the extinction process that they masked at the end of the extinction session. Although such effects may contribute to the deterioration of performance during the administration of an extinction procedure, they may have little to do with any underlying learning process. If spontaneous recovery could be accounted for solely on the basis of such effects, it would be of substantially less interest.

Contributors of this sort to response suppression seem especially likely to affect the standard spontaneous recovery designs that compare responding to the same stimulus in two different sessions. They are less likely to contribute to recovery when it is measured as the difference in responding to two stimuli during the same test session. Indeed, part of the motivation for developing the alternative assessment of spontaneous recovery (D) was to reduce the various general differences that are confounded when responding in two different sessions is compared.

Spontaneous Recovery Despite Loss of Acquisition Learning

Despite the widespread agreement that spontaneous recovery indicates that the primary extinction process does not involve the erasure of the original learning, there have been some attempts to preserve an erasure theory and still explain recovery. Perhaps the most well known attempt is the suggestion by Skinner (1938) that an extended extinction session leaves the stimuli from the beginning of the session incompletely extinguished. Consequently, those stimuli still have some strength that can be exhibited at the beginning of the next session. Skinner seems to have believed that extinction eventually removes all of the original learning but initially leaves the session-beginning stimuli partially effective, thereby explaining spontaneous recovery. Although there is some evidence that stimuli that are explicitly presented by the experimenter at the start of a session can retain some strength when extinction takes place in their absence (see Burstein and Moeser 1971), it seems unlikely that this will fully account for spontaneous recovery. Several experiments have found evidence of strong spontaneous recovery even when the session starting stimuli are well controlled (see Thomas and Sherman 1986) or when an extinguished stimulus is tested in the middle of a session only after another stimulus has completely lost its spontaneous recovery (see Robbins 1990). Moreover, it is hard to see why this account would anticipate one of the basic findings of spontaneous recovery, that it increases with the passage of time.

A somewhat more successful version of such an account was offered by Estes (1955) as a component of his influential stimulus sampling theory. Estes argued that what experimenters identify as stimuli can actually be viewed as constructed of many elements with occurrence that fluctuates in time. Acquisition and extinction produce changes in the associations between the out-

---

**Figure 2** An illustration of spontaneous recovery using the design shown in Figure 1D. Rat subjects were given Pavlovian magazine approach training and extinction with two stimuli, S1 and S2, and then tested in a common session. The superior responding in S1, compared with S2, is used to infer spontaneous recovery.
comes and those elements over the course of trials. According to that theory, when an element is paired with a reinforcer, it immediately becomes fully conditioned; when it is presented without reinforcement, it immediately reverts to its unconditioned state. It is only the random fluctuation in the selection of elements over trials that gives the normally observed gradualness to changes in behavior. At heart, this is a theory in which extinction involves the actual erasure of learning. The important point to notice is that when trials are given in close proximity, they tend to share more elements; when trials are separated, then they sample different elements from the pool that constitutes a given stimulus. Consequently, with repeated extinction trials, performance can decrease even though some unsampled elements remain conditioned. With the passage of time, the likelihood of sampling those nonextinguished elements can increase, generating spontaneous recovery.

Like the account offered by Skinner, this account correctly predicts that repeated extinction will diminish spontaneous recovery. Unlike that account, it gives no special role to session-initiating stimuli, and it can account for the growth in extinction with time. The stimulus sampling theory also correctly anticipates the occurrence of spontaneous recovery with all of the various assessment procedures. These successes demonstrate that one can account for at least some of the basic facts about spontaneous recovery even on the assumption that the extinction process involves (partial) removal of the original learning. They also highlight the fact that the observation of spontaneous recovery does not imply that extinction must leave all of original acquisition in place.

**Spontaneous Recovery as an Indication That Extinction Involves Inhibition**

If one envisions an extinction procedure as leaving intact even a portion of the original learning, then it is natural to conclude that some new learning is occurring that is superimposed on acquisition and prevents performance. The classic candidate for such new learning has been “inhibition.” A wide variety of different kinds of inhibition have been proposed, suggesting quite different ideas about the nature of extinction. But in each case, spontaneous recovery is accounted for by the dissipation of that inhibition with the passage of time.

**CS-Based Inhibition**

Pavlov (1927) was the first to suggest that extinction involves inhibition and that spontaneous recovery is an indication that this was so. He viewed the learned inhibition that he supposed to occur during extinction as more vulnerable than the excitation that develops in acquisition to such intrusions as the presentation of a new stimulus (i.e., disinhibition) and the passage of time (i.e., spontaneous recovery). Although it is not widely appreciated, Pavlov’s notion of inhibition was highly focused on the CS and was envisioned to develop independently of whether or not the reinforcer occurred. Pavlov believed that every CS presentation leads to the building up of a kind of fatigue in the neural cells stimulated by that CS, whether or not the reinforcer followed. This leads to a reduced ability of the CS to stimulate its neural targets, which would in turn lead to reduced behavior. That is, the natural consequence of repeatedly presenting the CS is a deterioration of behavior. However, during acquisition, this process is overwhelmed by the growth of an association which the CS develops with the reinforcer. With rest, the CS-based inhibition fades and responding can recover.

Robbins (1990) proposed a related notion of inhibition, which he described in terms of reduced attention to the CS. He suggested that during acquisition, attention to the CS grows whereas during extinction it falls. Part of the reason for the decline of behavior in extinction is then reduced attention to the CS with the attendant loss in its ability to evoke responding. With rest, that attention partially returns. Unlike Pavlov, Robbins envisioned the effectiveness of the CS as dependent on the trial consequence, growing with reinforcement and declining with nonreinforcement. In support of this view, Robbins (1990) found evidence, in a sign-tracking preparation with pigeons, that a CS lost its ability to serve a variety of different functions during an extinction procedure but then recovered them with time. He trained a CS simultaneously to have an excitatory association with a US and also to serve as a conditional signal that another CS would not be followed by a US. When he extinguished the excitatory association, the CS lost both properties and then regained them both with the passage of time. That result is consistent with the view that one contributor to the loss of behavior with extinction and its restoration with time may be changes in the processing of the CS per se. On the other hand, both Bouton and Peck (1992) and Rescorla (1997a) have found evidence for spontaneous recovery following counter-conditioning in which one US is replaced with another. Because counter-conditioning yields a CS that continues to produce some behavior, those cases of spontaneous recovery seem unlikely to be attributable to loss and recovery of the processing of the CS. Of course, the processes underlying spontaneous recovery after such procedures may differ from those underlying the changes after extinction.

**Response-Based Inhibition**

Shortly after Pavlov’s work became available in this country, Hull (1943) suggested a similar account of extinction and spontaneous recovery that focused on the response rather than on the stimulus. As part of a multiprocess account of extinction, Hull argued that each occurrence of a response leads to the building up of a fatigue-like process that is specific to that response and that reduces its subsequent likelihood of occurrence. This process contributes to the decrease in performance during extinction but dissipates with time, permitting spontaneous recovery. Similar to Pavlov, Hull seems to have believed that the changes in this response-fatigue are independent of reinforcement contingencies; indeed, he believed it left no permanent learning impact at all. That is, this particular inhibitory notion is more akin to the performance effects discussed above. It has proven difficult to confirm the most obvious implication of such an approach, that the effortfulness of the response should influence the course of extinction and spontaneous recovery (see Mackintosh 1974). But one advantage that has been claimed for Hull’s account is that spontaneous recovery described above is that it provides a common test session in which all stimuli should share any general fatigue processes.

**Outcome-Based Inhibition**

One way of viewing extinction is that a CS that was once followed by an effective outcome is now followed by an ineffective one. Indeed, there is evidence that if the outcome potency is deliberately partially reduced, then behavior partially deteriorates (see Wagner and Rescorla 1972; Kehoe and White 2002). This suggests the possibility that one change that occurs in extinction involves a reduction in the memory of the outcome itself. Based on this kind of thinking, Rescorla and his collaborators (see Rescorla and Heth 1975; Rescorla and Cunningham 1978) suggested that one process that occurs in an extinction procedure is a depression in what they called the “US representation.” This led them to explore manipulations that might affect the state of the US representation independently of its association with a particular CS. By using a fear conditioning preparation in rats, they documented the phenomenon of “reinstatement” first reported by Pavlov, that responding could be restored to an extinguished CS by separate presentations of the US alone.
Although others (see Bouton 1984) have given different accounts, Rescorla and Cunningham described this manipulation as restoring a portion of the extinction-depressed US representation. They also argued that there might be a recovery in the US representation with time, leading to the phenomenon of spontaneous recovery of responding to the CS. In support of that possibility they found that, under some circumstances, recovery could be undermined by the nonreinforcement of another CS immediately prior to testing, a manipulation presumed to reduce the US representation. However, various other implications of this notion have not received support. For instance, Robbins (1990) found evidence of independence in spontaneous recovery for two stimuli that had been conditioned with the same US. Moreover, Tomie et al. (1980) reported spontaneous recovery after an extinction procedure in which the US was continued but made independent of the CS, a procedure that should maintain the US representation. Nevertheless, it is possible that changes in the memory for the US may make a contribution to spontaneous recovery in some preparations.

**Associative Inhibition**

Most contemporary views of Pavlovian inhibition involve not individual events but rather associations between events. For instance, Konorski (1948) argued that a variety of conditioning paradigms, including extinction, lead to the development of an inhibitory association between the CS and US that is parallel to, but the opposite of, the excitatory associations set up during acquisition. In later writings, Konorski (1967) offered a somewhat different account, according to which inhibitory learning involves associations between the CS and a consequence that is the opposite of the US, the "no US." Most contemporary discussions of Pavlovian inhibition implicitly accept one or the other of these views.

The standard ways of detecting such associative inhibition are to ask whether an inhibitory stimulus can reduce responding to an otherwise effective excitor (a summation test) or is slow to acquire excitation if it now receives an excitatory training treatment (a retardation test; Rescorla 1969). Most contemporary evidence suggests that an extinguished stimulus does not in fact demonstrate net inhibition with either of these tests. Presumably this is because the inhibition that builds up during nonreinforcement is only just sufficient to counteract the original excitation, but does not develop beyond that level so as to generate a net inhibitory stimulus. However, there is some evidence that an originally neutral stimulus that accompanies an excitatory CS during an extinction procedure does capture inhibition, as assessed by these tests (see Rescorla 1979, 1999). Moreover, there are some reports (see Henderson 1978) that a net inhibitory stimulus does lose some of its inhibitory properties with the passage of time, in the way required if this were to contribute to spontaneous recovery.

An alternative associative basis for extinction has been proposed by Colwill (1991) and explored by Rescorla (1993). According to that account, the inhibition that develops during extinction involves associations between the stimulus and a response. These inhibitory S-R associations are envisioned as existing side by side with excitatory associations that involve the US. Consistent with this view (but not with the notion of inhibition between the CS and US), Rescorla (2001b) used outcome-specific transfer procedures to reveal that the CS-US associations maintained their full net strength after extinction, despite the reduction in responding. Moreover, he found no evidence that those CS-US associations changed in the way expected with the passage of time if they were to be the basis of spontaneous recovery. He proposed instead that extinction can involve an associative response-specific process that depresses responding but deteriorates with time.

An alternative view of Pavlovian conditioned inhibition in terms of modulation has received substantial recent attention. Several investigators (see Holland 1983, Rescorla 1985) have suggested that under some circumstances an inhibitory stimulus does not develop an association with a separate event, such as US or a response, but rather acquires the ability to control the strength of an excitatory CS-US association. This type of inhibition is not the opposite of an excitatory association but rather plays more of a hierarchical role, modulating a CS-US association. Bouton (1991) has suggested that both contextual and temporal stimuli can serve this modulatory role for a stimulus subjected to an extinction procedure. During the extinction procedure, the excitatory association is seen as remaining intact but being disabled by the presence of contextual or temporal stimuli that had accompanied nonreinforcement. With a change in either the context or the temporal stimuli, this inhibitory process is attenuated, allowing the return of performance.

The phenomenon of "renewal" provides evidence for this view. If, following extinction, the stimulus is tested in a different context, there can be a substantial recovery of responding (see Bouton 1991). If one views the passage of time as analogous to changing the context, this view provides a way of conceptualizing spontaneous recovery.

**Differential Retrieval**

The inhibitory accounts of extinction and spontaneous recovery all envision the strength of original learning as remaining unchanged throughout extinction and spontaneous recovery. They see the decrease in responding during extinction as attributable to the growth of the inhibitory process, and the increase in responding with spontaneous recovery as attributable to the dissipation of that inhibitory process. In effect, the memory for the extinction process loses strength with the passage of time.

An alternative framework for understanding extinction and spontaneous recovery involves differential retrieval of the acquisition and extinction experiences, both of which remain fully intact. For instance, Spear (1971) and Bouton (1993) have both noted the formal analogy between interference paradigms as studied in humans and the acquisition/extinction sequence of Pavlovian conditioning. In both cases, the organism is exposed to two competing pieces of information that might be expected to interfere with each other. This led both investigators to the proposal that differential performance might reflect differential likelihood of retrieving the two experiences, rather than a weakening of stored information about either experience. According to this view, manipulations that make the retrieval of one experience or the other more likely will result in changes in overall performance. For instance, a stimulus that accompanies extinction might serve as a retrieval cue, presentation of which would promote the retrieval of the memory for extinction. Evidence for such a process can be found in the ability of some such stimuli to diminish the magnitude of spontaneous recovery (see Brooks and Bouton 1993).

A related view, focusing specifically on spontaneous recovery, is incorporated in Davenport’s (1998) temporal weighting rule. According to that rule, when an animal has multiple experiences with a stimulus prior to a test, it weights those experiences according to the relative time that has passed between each and the test. In the case of acquisition followed by extinction, this means that performance will deteriorate under the currently experienced nonreinforced treatment. However, as time passes, the relative temporal advantage enjoyed by the recent nonreinforcement experience will diminish, leading to spontaneous recovery. On views such as these, there is a permanence for both
the acquisition and extinction experiences; what changes with time is their relative likelihood of retrieval.

As this discussion indicates, a broad range of different processes have been envisioned as contributing to the decrement in performance resulting from an extinction procedure and to the recovery from that decrement with the passage of time. Moreover, one can identify evidence supporting each idea. It seems likely that each of these may contribute to the changes in various situations. However, none of them seems sufficient to provide an account on its own.

One challenge that each contributor faces is to provide an account of the demonstrated empirical determinants of spontaneous recovery. For this reason it is worth reviewing some of those determinants.

**Basic Empirical Properties**

Despite the centrality of spontaneous recovery to the phenomenon of extinction, there is actually surprisingly little well-documented information on its detailed properties. However, four features of spontaneous recovery seem to be widely accepted.

**Spontaneous Recovery Increases in a Negatively Accelerated Fashion Over Time**

Almost every description of spontaneous recovery includes the claim that recovery is greater the greater the delay between extinction and test. Indeed, there is a sense in which greater recovery with more time is a defining property. Moreover, the form of that increase appears to be negatively accelerated. For instance, in recent years, negative accelerated recovery has been found with eyelid conditioning in rabbits (Haberlandt et al. 1978), sign-tracking in pigeons (Robbins 1990), and fear-conditioning in rats (Quirk 2002). Almost all of the potential contributors to recovery listed above appear to be consistent with such a pattern of change. Although the various inhibitory theories make no specific predictions about how that inhibition fades with time, the negatively accelerated nature of many biological processes makes the finding unsurprising. The stimulus sampling mechanism described by Estes (1955) and the differential weighting rule proposed by Devenport (1998) also yield this expectation. Perhaps the only potential contributor that does not anticipate this is Skinner’s (1938) suggestion that at the beginning of the session stimuli retain their ability to evoke a response.

**Spontaneous Recovery Is Incomplete**

Although Pavlov claimed to have observed full recovery from extinction, most other investigators have reported only partial recovery. Even when recovery of responding appears to be complete on the first test trial, the rapid loss of responding over the course of testing suggests that recovery was actually less than full (notice, for instance, the rapid loss during testing shown in Fig. 2). The incompleteness of recovery appears to be mandated by some of the mechanisms described above. For instance, the stimulus sampling account of Estes and the weighting rule of Devenport appear incapable of allowing the impact of extinction to disappear altogether. The other mechanisms are less well-specified in this regard.

**Spontaneous Recovery Declines With Repeated Extinction**

It is widely agreed that the greater the amount of extinction, the less the magnitude of spontaneous recovery after any fixed waiting time. This most frequently shows up as the reduced amount of recovery from day to day over multiple days of extinction. An illustration of that decline is shown in Figure 3, which displays results from a recent magazine-approach study with rat subjects done in our laboratory. That figure shows responding during repeated daily extinction sessions in which a 30-sec noise was presented eight times without its food pellet. Another stimulus, a light, received intervening reinforced trials. It is clear that there is repeated recovery of responding that gets smaller over the course of extinction.

Actually, in the absence of some better understanding of scaling issues, it is difficult to compare quantitatively the amounts of recovery for stimuli that have undergone different amounts of extinction. Clearly, if minimal extinction has occurred, there is less decrease in behavior and hence less opportunity for recovery to occur. Similarly, with massive extinction, performance may be at a floor, and hence, even substantial recovery may be difficult to detect. But most mechanisms of recovery appear to anticipate that the deeper the extinction the less the recovery. That prediction is obligatory for accounts of recovery such as that offered by Estes and by Devenport, who see repeated extinction as accumulating a relatively permanent depressive process. Data such as those shown in Figure 3 are commonly taken as consistent with this prediction.

**Spontaneous Recovery Can Be Brought Under Stimulus Control**

There is good evidence that the learning process that occurs during extinction can be brought under stimulus control. For instance, Rescorla (1979) reported that a stimulus present during the nonreinforcement of a previously trained excitor took on the properties of a conditioned inhibitor, suppressing responding to other exciters. Indeed, the so-called conditioned inhibition paradigm consists of little more than intermixing reinforced and nonreinforced presentations of a stimulus while signaling the nonreinforced presentations by a second stimulus. Similarly, Bouton (1991) has reported that if the context present during extinction is removed by testing in another context, then the suppressive effects of extinction are reduced and behavior is “renewed.” Brooks and Bouton (1993) have extended these observations to the responding observed in spontaneous recovery. They found (see also, Brooks 2000) that if an explicit stimulus is present during extinction of an excitor, then that stimulus has the ability to diminish spontaneous recovery if it is presented at the time of the test.

Although few theories of extinction are challenged by the observation that whatever is occurring in extinction can be brought under the control of a stimulus, retrieval theories seem like the most natural account. For instance, Bouton has argued that a stimulus present during extinction is especially good at retrieving a memory for a CS–US association. There is now substantial evidence that one stimulus can be learned as a signal of the relation between another stimulus and the US (see Schmajuk and Holland 1998). Such modulation or “occasion setting” could
be the mechanism by which explicit stimuli, contexts, or even time, might activate the memory of a CS-US association (see Bouton 1991, 1993).

Determinants Needing Further Investigation
In addition to these four well-established findings, there are a variety of other manipulations that have been claimed to affect spontaneous recovery but for which there is substantially less or even contradictory evidence. This is unfortunate because the effects of many of these manipulations might be informative in identifying the contributions of particular mechanisms of recovery.

The Symmetry of Spontaneous Recovery and Regression
Accounts of spontaneous recovery differ in the degree to which they treat extinction as engaging a special learning process with distinctive properties, such as the likelihood of its memory fading in time. Beginning with Pavlov’s, the various proposals of inhibitory processes have tended to see them as different from excitatory process precisely in their greater instability with the passage of time. This is clearly true for the fatigue-like processes mentioned by Pavlov, Robbins, and Hull, but it also seems true of some associative inhibition accounts, such as those described by Rescorla and Bouton. By contrast, the stimulus sampling theory of Estes and accounts that appeal to retrieval or relative weighting seem to make little distinction between acquisition and extinction processes in their inherent vulnerability to time. They see the animal as integrating two experiences that it receives sequentially in time in a similar way regardless of the identity of those processes. This means that the latter accounts anticipate that one should observe a companion phenomenon to spontaneous recovery from extinction if one were to interchange the order in which extinction and acquisition were administered. That is, they expect to see regression of responding after acquisition if that acquisition were preceded by some sort of nonreinforced training.

The evidence for such regression is highly mixed. Notice that the simple deterioration of performance from day to day during acquisition is not sufficient to identify regression that is the opposite of spontaneous recovery. The critical observation is that there is a deterioration in performance that is attributable to a prior history of nonreinforcement, just as the critical observation for spontaneous recovery is that there is an improvement with time that is specific to stimuli that have a history of reinforcement.

When animals are given in sequence two reinforcement experiences that differ in reinforcer valence or frequency, there is evidence that regression in the direction of the first performance can be observed with time (see Bouton and Peck 1992; Mazur 1996). But the results are less clear when nonreinforcement of a stimulus precedes reinforcement prior to the waiting period. Some early experiments reported positive results (see Spear et al. 1965; Konorski 1967). But some more recent studies have found no evidence for regression or the opposite results (see Kraemer et al. 1991; Rescorla 2001b). Clearly it would be valuable to understand the circumstances under which one obtains either regression or its opposite if one is to evaluate the contributions of various mechanisms to spontaneous recovery. It will surely be important to distinguish among different kinds of nonreinforcement experience that might precede reinforcement. The simple exposure to a stimulus prior to any reinforcement is certain to endow it with properties different from those of a stimulus that signals nonreinforcement explicitly, as in the case of conditioned inhibition training or even extinction. But there are not sufficient data to indicate whether or not this distinction matters for the production of regression.

Recovery Following Massed or Spaced Extinction Trials
There is reason to anticipate that conducting extinction with short intertrial intervals may encourage more rapid response decrement followed by more substantial recovery. Certainly this is the expectation of accounts such as that of Hull, which emphasizes short-term fatigue-like effects, and of Estes, which emphasizes that massing of trials would yield repeated sampling of the same stimulus elements but neglect of others. Indeed, one might argue that there is a logical sense in which spaced trials should lead to slower decrement and less recovery. Presumably widely spaced trials would allow for any recovery between individual trials, resulting in slower behavioral loss over the course of an extinction procedure but more substantial change by the time that a test for recovery is imposed.

Despite the appeal of these arguments, the evidence on the impact of massing or spacing extinction trials is quite mixed. A number of investigators (see Rescorla and Durlach 1987; Cain et al. 2003) have reported that massing produces rapid loss of performance. However, Rescorla and Durlach reported no difference in the magnitude of responding in a subsequent test for spontaneous recovery and Cain et al. (2003) reported continued less responding after massed extinction even with the passage of time. To complicate matters further, Stanley (1952) reported that for an instrumental training task, massing slowed extinction on one measure and speeded it on another in an instrumental choice situation.

Interval Between Learning and Extinction
Although most attention has focused on the interval between extinction and the recovery test, it is also of interest to ask about the impact of the interval between the original training and extinction, as a determinant of spontaneous recovery. The retrieval theory proposed by Spear and the weighting rule described by Davenport both suggest that spontaneous recovery should be maximal when the interval between acquisition and extinction is minimized. In both cases, the intuition is that when training and extinction are close in time, it should be more difficult for the animal to recall which is the more recent. Immediately after extinction, the relative temporal recency of the nonreinforced experience should be maximal. However, as time passes, and the two experiences are more equally distant in time, they should become more equivalent in their likelihood of being retrieved. The increase in the relative likelihood of retrieving the original acquisition experience would then produce spontaneous recovery. A similar reasoning would lead to the relatively greater impact of the acquisition experience according to the weighting rule. Mechanisms of recovery that appeal to the loss of the extinction experience have no natural way to predict that the interval between training and extinction should matter.

Unfortunately, there are very few attempts to examine this possibility. There is some supportive evidence from studies of proactive inhibition in humans (Underwood and Freund 1968) and from counter-conditioning in rats (Gordon and Spear 1973), but little for the case of extinction. Recently, Rescorla (2004) has reported that a longer time interval between training and test diminishes spontaneous recovery in magazine approach and instrumental responding in rats and in sign-tracking in pigeons. One illustration is shown in Figure 4. That figure displays the results of extinction and testing with two stimuli given acquisition, extinction, and a test for spontaneous recovery in a magazine-approach procedure using rat subjects. The stimuli shared the same 48-h recovery interval after extinction but differed in that a greater interval (8 d versus 1 d) intervened between training and extinction for S1 than for S2. The two stimuli showed virtually identical behavior over the course of extinction. However, the results of a test for spontaneous recovery given 48 h...
after extinction show greater spontaneous recovery for the stimulus with the shorter training/extinction interval (S2).

Results such as these suggest that, consistent with some retrieval theories, spontaneous recovery may be a decreasing function of the acquisition/extinction interval. But clearly more work needs to be done on this potentially informative parameter.

Conclusion

The picture that emerges from this discussion of spontaneous recovery is one of a process that is greatly in need of further empirical investigation. The available evidence fails to identify any one proposed process as the sole basis for spontaneous recovery. However, there is also evidence in support of all of the processes occurring in extinction, one cannot remove all of original acquisition seems secure. Spontaneous recovery is joined by a variety of other phenomena, such as disinhibition, renewal, reinstatement, and augmented summation (see Reberg 1972) as a basis for that inference. But the simple observation of spontaneous recovery does not force the inference that all of original learning remains nor even that the learning that occurred during extinction fades in time.

In the light of this conclusion, it is unfortunate that we do not have a clearer picture of how some of the parameters of most potential interest affect spontaneous recovery. But it suggests that if one is to use spontaneous recovery as a tool to understand the nature of the processes occurring in extinction, one cannot simply celebrate its occurrence or its failure to occur. We will have to do much more analytic experiments determining the circumstances under which it occurs in the particular extinction situation under study.

REFERENCES

Amsel, A. 1958. The role of frustrative nonreward in noncontiguous reward situations. Psychol. Bull. 56: 102–119.

Bouton, M.E. 1991. Context and retrieval in extinction and in other examples of interference in simple associative learning. In Current topics in animal learning: Brain, emotion, and cognition (eds. L. Dachowski and F. Flaherty), pp. 25–53. Lawrence Erlbaum, Hillsdale, NJ.

Bouton, M.E. and Peck, C.A. 1992. Spontaneous recovery in cross-modal transfer (counterconditioning). Anim. Learn. Behav. 20: 313–321.

Brooks, D.C. 2000. Recent and remote extinction cues reduce spontaneous recovery. J. Exp. Psychol. Comp. Physiol. Psychol. 53B: 25–58.

Brooks, D.C. and Bouton, M.E. 1993. A retrieval cue for extinction attenuates spontaneous recovery. J. Exp. Psychol. Anim. Behav. Processes 19: 77–89.

Burstein, K.R. and Moeser, S. 1971. The informational value of a distinctive stimulus associated with the initiation of acquisition trials. Learn. Motivat. 2: 228–234.

Cain, C.K., Blouin, A.M., and Barad, M. 2003. J. Exp. Psychol. Anim. Behav. Processes 29: 323–333.

Colwill, R.M. 1991. Negative discriminative stimuli provide information about the identity of omitted response-contingent outcomes. Anim. Learn. Behav. 19: 326–336.

Delamater, A.R. 2004. Experimental extinction in Pavlovian conditioning: Behavioural and neuroscience perspectives. Q. J. Exp. Psychol. 57B: 97–132.

Devenson, L.D. 1998. Spontaneous recovery without interference: Why remembering is adaptive. Anim. Learn. Behav. 26: 172–181.

Estes, W.K. 1955. Statistical theory of spontaneous recovery and regression. Psychol. Rev. 62: 145–154.

Gordon, W.C. and Spear, N.E. 1973. Effect of reactivation of a previously acquired memory on the interaction between memories in the rat. J. Exp. Psychol. Anim. Learn. Motivat. 9: 149–155.

Haberlandt, K., Hamsher, K., and Kennedy, A.W. 1978. Spontaneous recovery in rabbit eyelid conditioning. J. Gen. Psychol. 98: 241–244.

Henderson, R.W. 1978. Forgetting of conditioned fear inhibition. Learn. Motivat. 9: 16–30.

Holland, P.C. 1983. “Occasion-setting” in Pavlovian feature positive discriminations. In Quantitative analyses of behavior: Discrimination processes, Vol. 4 (eds. M.L. Commons et al.), pp. 183–206. Ballinger, Cambridge, MA.

Hull, C.L. 1943. Principles of behavior. Appleton-Century-Crofts, New York.

Kehoe, E.J. and White, N.E. 2002. Extinction revisited: Similarities between extinction and reductions in US intensity in classical conditioning of the rabbit’s nictitating membrane response. Anim. Learn. Behav. 30: 96–111.

Konorski, J. 1948. Conditioned reflexes and neuron organization. Cambridge University Press, Cambridge, UK.

Kraemer, P.J., Randall, C.K., and Carbary, T.J. 1991. Release from latent inhibition with delayed testing. Anim. Learn. Behav. 19: 139–145.

Mackintosh, N.J. 1974. The psychology of animal learning. Academic Press, New York.

Mazur, J.E. 1996. Past experience, recency, and spontaneous recovery in choice behavior. Anim. Learn. Behav. 24: 1–10.

McSweeney, F.K. and Swindell, S. 1999. General-process theories of motivation revisited: The role of habituation. Psychol. Bull. 125: 437–457.

Myers, K.M. and Davis, M. 2002. Behavioral and neural analysis of extinction. Neuron 36: 563–584.

Quirk, G.J. 2002. Memory for extinction of conditioned fear is long-lasting and persists following spontaneous recovery. Learn. Memory 9: 402–407.

Pavlov, I.P. 1927. Conditioned reflexes. Oxford University Press, Oxford.

Reberg, D. 1972. Compound tests for extinction. Anim. Learn. Behav. 20: 917–920.

Rescorla, R.A. 1969. Pavlovian conditioned inhibition. Psychol. Bull. 72: 77–94.

Rescorla, R.A. 1979. Conditioned inhibition and extinction. In Associative mechanisms in conditioning: A memorial volume for Jerzy Konorski (eds. R.A. Boaakes and A. Dickinson), pp. 83–110. Erlbaum, Hillsdale, NJ.

Rescorla, R.A. 1985. Conditioned inhibition and facilitation. In Information processing in animals: Conditioned inhibition (eds. R.R. Miller and N.S. Spear), pp. 299–326. Erlbaum, Hillsdale, NJ.

Rescorla, R.A. 1988. Behavioral studies of Pavlovian conditioning. Ann. Rev. Neurosci. 11: 329–352.

Rescorla, R.A. 1993. Inhibitory associations between S and R in extinction. Anim. Learn. Behav. 21: 327–336.

Rescorla, R.A. 1997a. Spontaneous recovery after Pavlovian conditioning with multiple outcomes. Anim. Learn. Behav. 25: 99–107.

Rescorla, R.A. 1997b. Spontaneous recovery of instrumental discriminative responding. Anim. Learn. Behav. 25: 485–497.
Spontaneous Recovery

Rescorla, R.A. and Cunningham, C.I. 1978. Recovery of the US representation over time during extinction. Learn. Motivat. 9: 373–391.
Rescorla, R.A. and Durlach, P.J. 1987. The role of context in intertrial interval effects in autoshaping. Q. J. Exp. Psychol. 39B: 35–48.
Rescorla, R.A. and Heth, C.D. 1975. Reinstatement of fear to an extinguished conditioned stimulus. J. Exp. Psychol. Anim. Behav. Processes 1: 88–96.
Rescorla, R.A. and Holland, P.C. 1976. Some behavioral approaches to the study of learning. In Neural mechanisms of learning and memory (eds. E. Bennett and M.R. Rosenzweig), pp. 165–192. MIT Press, Cambridge.
Robbins, S.J. 1990. Mechanisms underlying spontaneous recovery in autoshaping. J. Exp. Psychol. Anim. Behav. Processes 16: 235–249.
Schmajuk, N.A. and Holland, P.C., eds. 1998. Occasion setting: Associative learning and cognition in animals. American Psychological Association, Washington, DC.
Skinner, B.F. 1938. The behavior of organisms. Appleton-Century-Croft, New York.
Spear, N.E. 1971. Forgetting as retrieval failure. In Animal memory (eds. W.K. Honig and P.H.R. James), pp. 45–109. Academic Press, New York.
Spear, N.E., Hill, W.F., and O'Sullivan, D.J. 1965. Acquisition and extinction after initial trials without reward. J. Exp. Psychol. 69: 25–29.
Stanley, W.C. 1952. Extinction as a function of the spacing of extinction trials. J. Exp. Psychol. 43: 249–260.
Underwood, B.J. and Freund, J.S. 1968. Effects of temporal separation of two tasks in proactive inhibition. J. Exp. Psychol. 76: 50–54.
Wagner, A.R. and Rescorla, R.A. 1972. Inhibition in Pavlovian conditioning: Application of a theory. In Inhibition and learning (eds. R.A. Boakes and S. Halliday), pp. 301–356. Academic Press, New York.
Spontaneous Recovery

Robert A. Rescorla

Learn. Mem. 2004, 11:
Access the most recent version at doi:10.1101/lm.77504

References
This article cites 33 articles, 1 of which can be accessed free at:
http://learnmem.cshlp.org/content/11/5/501.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.