On the Rate of Extinction of a Conditioned Reflex

B. F. Skinner

Laboratory of General Physiology of Harvard University, USA

Published online: 06 Jul 2010.

To cite this article: B. F. Skinner (1933) On the Rate of Extinction of a Conditioned Reflex, The Journal of General Psychology, 8:1, 114-129, DOI: 10.1080/00221309.1933.9713178

To link to this article: http://dx.doi.org/10.1080/00221309.1933.9713178

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly...
ON THE RATE OF EXTINCTION OF A CONDITIONED REFLEX*

From the Laboratory of General Physiology of Harvard University

B. F. Skinner

I

A method of investigating the rate of formation of a conditioned reflex has been described in a previous paper (1). Without any essential modification it is suitable also for the study of experimental extinction, which differs from conditioning only in the direction of the observed change. In conditioning, an originally weak "investigatory" reflex, arbitrarily chosen for the experiment, is strengthened when the elicitation of the response is followed immediately by a second stimulus leading to the ingestion of food. In extinction, on the other hand, the response, once conditioned, is then not followed by the second stimulus, and the strength of the reflex diminishes. Thus conditioning appears experimentally as an increase in reflex strength, extinction as a decrease. In both cases the rate at which the response is elicited under constant stimulation may be used as a measure of the strength. The process of extinction is investigated in the following way.

The general procedure has been described in reporting the experiments on conditioning (1). The apparatus consists, in part, of a food-magazine, a food-tray, and a lever, so arranged that when the lever is pressed downward (by a rat) the magazine discharges a small pellet of a standard food into the tray. Four sets of the apparatus are operated simultaneously. In investigating extinction the response to the lever is first maximally conditioned. The connection between the magazine and lever is then broken, and the subsequent responses of the rat are recorded in the usual manner.1

*Accepted for publication by W. J. Crozier of the Editorial Board.
1As an alternative method the magazine may be left functional but empty. The response to the lever is then followed as usual by the click of the magazine. Responding to the click, the rat approaches the food-tray, but the stimulation ordinarily supplied by the pellet of food is lacking, and the chain is broken at that point. Although this is a relatively complex case, it seems to involve no significant difference in the rate or in the change in rate of extinction. A minor difference in the behavior is mentioned in Footnote 5.
panying figures are reproduced from kymograph records, to which coordinates have been added. Each response is represented by a short (uniform) elevation in the recorded curve. The rate of elicitation is given by the slope, and the records are, in effect, number-time graphs of the responses to the lever. As a necessary modification in dealing with the higher rates observed in extinction, the kymographs are here moving twice as fast as in the experiments reported in (1).

II

Four male rats, members of the 18th generation of an inbred strain,² 98 days old, were conditioned in the manner previously de-

²Known in this laboratory as the P strain. For its genetic history see Crozier, W. J., and Pincus, G., J. Gen. Physiol., 1931, 15, 243.
scribed. Records of the conditioning of this group will appear in a later communication. At the time of the conditioning the rats were permitted to obtain about 100 pellets of food, pressing the lever once for each pellet. They were then removed from the apparatus, and extra food was given them for one hour. On the following day the magazines were disconnected, so that the movement of the lever was without effect, except upon the recording devices in an adjoining room. The rats were placed in the experimental boxes and released at the usual time and in the usual way. Their responses to the lever during the following hour are recorded in Figure 1. The animals were released approximately at the beginning of the curves, except Rat D, which showed a characteristic delay in going through the one-way door of the release compartment, and whose record begins about five minutes after the release.

When the first response to the lever fails to supply the stimulus for the next member of the usual chain, the response is elicited again immediately, and a high rate of elicitation is maintained for a short time. It is evident from the figure, however, that this is soon interrupted and that the rate subsequently undergoes an extensive fluctuation. In spite of the irregularity it is possible to indicate a general course for each curve. When this is done, as in Figure 1, it is apparent that the deviations have the character of depressions in the curve, and that they are followed by compensatory increases in rate. This point is discussed in detail later. Because of the compensation the smooth curve is eventually fairly closely approximated. The theoretical curves in Figure 1 are logarithmic. Their agreement with the experimental curves is also discussed later.

At the end of the experimental period represented in Figure 1 a low rate of elicitation has been reached. The strength of the reflex has been reduced approximately to its value prior to conditioning. But if the rats are returned to the apparatus on the following day (no reconditioning having taken place), small extinction curves are obtained. A "loss of extinction" has occurred in the intervening 23 hours. Such an effect ultimately disappears if the extinction is sufficiently prolonged, and it is probable that a state of complete extinction can be reached. Since the reflex may persist, however, as an investigatory reflex in spite of the extinction, we are limited to the above statement that the frequency of elicitation has been reduced approximately to its value prior to conditioning.

When any degree of extinction has been attained, the reflex may
FIGURE 2

Record of reconditioning and subsequent extinction.

At A: some loss of the very thorough extinction obtained on the preceding day. At B: reconditioning as ten pellets of food are obtained. At C:

See text.
be restored to full strength by reconditioning. It may then be extinguished again in the same way, subject to the modification noted below. In Figure 2, taken with another rat, the reflex had been thoroughly extinguished on the preceding day, with the alternative method described in Footnote 1. Some "loss of extinction" is shown in the small curve at A. The rate then reached a very low value. At B 10 pellets of food were put into the magazine, and these were obtained by the rat with the 10 responses shown at that point. The resulting extinction follows at C. A base-line for the curve for extinction has been added. The details of the curve are discussed later.

That such an increase in strength is actually a reconditioning and not simply a facilitative effect due to the reception of food is shown in Figure 3. In this experiment the reflex had been very thoroughly extinguished. There is a slight loss of extinction evident at the beginning of the curve, followed by only two responses during the next 25 minutes. Three pellets of food were then discharged from the magazine by the experimenter, the lever remaining disconnected. The rat pressed the lever only five times in the following 18 minutes, indicating that a facilitative effect, if present, was very slight. At the next response the magazine had been connected, and the rat obtained three pellets of food, pressing the lever once for each. The magazine was then disconnected and a significant curve for extinction obtained.

At practically complete extinction the strength of the reflex is equal in value to its strength prior to conditioning, and there is no observable modification in the behavior of the rat. The process of extinction has thus simply reversed the change observed in conditioning. Nevertheless, the combined process of conditioning-and-

---

*The acceleration exhibited in the curve for reconditioning is not always obtained under apparently comparable conditions. If it is real, it may represent an inhibitory effect of the preceding extinction, but the point cannot be settled on the existing evidence.*
extinction has left a detectable effect upon the organism, which appears experimentally, not as a change in immediate behavior, but as a modification in the rate or in the rate of change of the rate of some subsequent process. When, for example, a reflex is reconditioned and then extinguished a second time (the two amounts of conditioning being roughly equal), the area under the second curve for extinction is reduced. The strength of the reflex, measured at the beginning of the two processes, may be the same, but the change in strength proceeds much more rapidly in the second case. Thus there is no simple relation between the strength of the reflex at a given moment and what might be called its resistance to extinction beyond that moment.

A similar effect, an example of which may be mentioned for its bearing upon extinction, is commonly known as over-conditioning. We have seen elsewhere that when the response to the lever is originally conditioned the reflex may reach an essentially maximal strength upon only one occasion when its elicitation is followed by the discharge of food from the magazine (1). It is probable, however, that a curve for extinction of the magnitude of those shown in Figure 1 could not be obtained after only one elicitation of this sort. The other hand, curves of greater magnitude might have been obtained in Figure 1 if the number of responses when food was obtained had been greater. Although there may be no significant increase in the immediately observed strength of the reflex after the first (conditioning) elicitation, there is possibly an extensive increase in resistance to extinction. Nevertheless, a high degree of resistance may be reached in a very short time, as shown in Figure 4. Here the first five responses indicate the practically instantaneous conditioning of the reflex. The initial response was the first in the history of the rat to be followed by the delivery of food. Although there was a slight delay after the second response, the average rate for the first five is very near the maximal for a normal eating curve. After the fifth response the magazine was disconnected and the reflex extinguished as shown. Although conditioning and over-conditioning were possible at only five elicitations, the curve for extinction is composed of more than two hundred responses, and the rate has not then closely approached zero. The details of the curve, especially the two major breaks at A and B, are discussed later.

The record thus confirms the observations reported in (1). In comparing the records the difference in the time coordinates should be noted.
The beginning of this record shows the original conditioning of the reflex. The first five responses (as far as the horizontal mark) are the first and only responses to the lever to be followed by the delivery of food. The subsequent extinction is shown in the rest of the record. The breaks at A and B are due to conflicting stimuli inadvertently introduced into the experimental box.
What we have called resistance to extinction is properly beyond the scope of this paper. It belongs to a class of effects which have so far failed to receive an adequate experimental treatment. They present a special problem, because, as we have noted, they appear as modifications, not in immediate behavior, but in the rate of change of the rate of some process. A quantitative investigation of "resistance to extinction" would presuppose some method of measuring the course of the extinction itself and some elementary knowledge of its nature. We are concerned here with these preliminary steps.

III

Sixteen records of original extinction and several hundred of extinction-after-reconditioning have been examined. None of them calls for any significant modification of the statements in the preceding section. The specimens reproduced are typical. In most cases better approximations to a smooth curve could have been given, but they would have represented the average result less accurately. In general, the curves show marked deviations; and this must be accepted as characteristic of the change in reflex strength observed under these conditions. The deviations are not, however, haphazard. The intervals elapsing between successive responses are distributed by no means at random, and the observed fluctuations are, in that sense, real. Before giving some account of their nature, it will be well to consider several possible sources of disturbance in the recorded curves and certain incidental changes in the behavior of the rat, which must be taken into account in determining the course of the major process.

There is, first of all, the possibility, elsewhere discussed, that the recording apparatus may fail to follow the behavior of the rat. The normal eating curve is regular enough to expose errors of this sort, but records obtained with ordinary care seem to be free of them. In the process of extinction, however, much higher rates must be dealt with. A greater strain is placed upon the apparatus and particularly upon the device introduced to meet the chief difficulty of the method, which arises from the use of an electrical contact in recording the displacement of the lever (3, 4). To check the accuracy of the apparatus a special signal has been used which operates independently of the recording device and enables the experimenter to compare the record with the actual make-and-break in the circuit. The present records are believed to be correct reports of the movement of the lever.
Secondly, it is possible that the character of the response may change during extinction, in such a way that the closing of the mercury switch at the lever will no longer serve as a reliable criterion of the behavior. But this is evidently not true. Throughout the course of the extinction the contacts are, in general, cleanly made and broken, and their duration holds rather closely to a constant average value.  

A third possible source of irregularity is the base line upon which the process of extinction is observed. An unconditioned rat will occasionally respond to the lever. A certain number of these "investigatory" responses will be recorded as the inevitable result of the activity of the rat within the narrow limits of the experimental box. The rate of elicitation is irregular and varies with the hunger of the animal. Responses of this sort offer no difficulty in recording a normal eating curve, because the ingestive reflexes are prepotent during the early part of the curve and the investigatory reflexes are weak during the latter. In the process of extinction, however, the rate of elicitation approaches, not zero, but a value characteristic for investigatory reflexes in a hungry rat. This is a real, if not highly important, factor. It means in practice that we may expect to find a fairly small number of responses, distributed at random throughout the curve, that appear as deviations from an otherwise rational course. [It is well to make a special class of those instances in which the displacement of the lever is the result, not of a response, but of the movement of the rat under the influence of mechanical forces. Such would be the case, for example, if the rat fell against the lever. Compare (2, p. 444).]

We must also consider certain incidental deviations from a smooth curve, correctly reported by the apparatus, which reflect genuine changes in the behavior of the rat. We have discussed an example elsewhere, in dealing with the modification in reflex strength observed during the ingestion of food (3,4). In the normal eating behavior of the rat short periods frequently occur in which no eating takes place. They are followed typically by compensatory increases
in rate, so that only temporary divergences from the normal eating curve result. It is easily shown that such an effect is not part of the process represented by the main curve. The delays occur haphazardly during an experimental period, and they may be absent altogether. They are more frequently observed when the reflex is relatively weak, and this suggests that they are due to the operation of conflicting, momentarily prepotent stimuli.

The same effect is observed when the response is not followed by the ingestion of food, as in the process of extinction. In the present curves, however, most of the examples are obscured by the rapid fluctuation in rate, into which, as we shall see shortly, the delays and recoveries have become incorporated. Since the effect is used in the following analysis, some independent evidence is called for. Fortunately, curves for extinction can be obtained that are not subject to the characteristic fluctuation observed with the present method. The necessary procedure will be described in a later communication, where examples of the effect in question will be given. A deviation will also be demonstrated, of a converse sort, in which a brief increase in rate is subsequently compensated for by a period of delay. The existence of both effects may for the present simply be assumed.6

IV

From a casual inspection of Figures 1, 2, and 4, it is apparent that the typical curve for extinction is wave-like in character. There is no very uniform wave-length or amplitude, but the rate of elicitation clearly tends to pass from a high to a low value and back again. The transitions are sometimes smoothly executed but are more often abrupt. Since an effect of this sort is presumably continuous, we may regard the smooth transition as representative of the actual change in strength, and we may then account for the abrupt transitions by appealing to some of the disturbing factors discussed in the foregoing section.

6An example of the effect of extraneous stimuli was obtained inadvertently in the experiment recorded in Figure 4, where a short piece of heavy twine was left projecting into the box and was chewed to bits by the rat during the course of the experiment. This must have been accomplished during the two major breaks at A and B, for here only could the rat have found the necessary time. The breaks are unusual (considering the slope at which they appear) and call for some such special explanation. Granted the present one, the record offers an excellent example of the effect of conflicting stimuli and of the compensating process of recovery.
If we assume a continuous change in strength as given by the curved lines in Figure 5, then the best possible representation with the present method is shown at A in that figure. The weak part of such a curve is the middle section, where the strength is low, and if a break is observed, it will probably be in that region. A delay of the sort just described, together with its subsequent recovery, would affect the record approximately as shown at B. The deviation may appear at some other part of the curve, however, and the result will then be otherwise, for example, as at C. The converse type of deviation will give other characteristic effects. If the acceleration appears, as it most often does, when the rate is passing from a high to a low value, the theoretical curve will be overshot as shown at D, and the latter half of the cycle may then be of either the A or the B type.

We have used for illustration a complete cycle composed of 16 responses and have assumed an incidental deviation affecting about one fourth that number. The ease with which a deviation can be identified and allowed for will depend, of course, upon this ratio. Consequently, it is important to notice that both the wavelength and the magnitude of the deviation may vary widely from these selected values. Furthermore we have assumed a simple and quite arbitrary wave-form. A quantitative treatment of the fluctuation would require a closer reference to the actual data than this elementary scheme possesses. Our present purpose, however, is to
show simply that the curves are wave-like in nature; and this can be done qualitatively by pointing out that the records in Figures 1, 2, and 4 can, in general, be divided into cycles corresponding to one or another of the types in Figure 5. The principal examples in the former figures have been lettered (lower case) according to type. The effects are not always clear-cut, but we have seen that the curves are subject to other minor disturbances. The probable course of the change in strength in the four curves in Figure 1, as distinct from the disturbed course given experimentally, may be represented as in Figure 6.

A simple theoretical account of this cyclic fluctuation can be derived from the assumption that an interruption of the normal chain
of ingestive reflexes sets up a depressive effect that lowers the rate of elicitation. In order to account for the present observations we must assume a time-lag for the effect and must appeal also to a previously demonstrated compensatory increase in rate following an incidental delay (3). In the process of extinction the rate of elicitation begins at a maximum. Under the conditions of the experiment each elicited chain is interrupted when the necessary stimuli for the continuance of the chain are not supplied, and, by assumption, a depressive effect is generated. A period of reduced rate supervenes, during which the previously generated effect disappears and (because of the lowered rate) is not replaced to any considerable extent. Through a compensatory increase the rate returns to a maximal value, where it enters upon a second cycle. The properties of the resulting curve will depend upon the extent of the depressive effect of a single interrupted elicitation, upon the time-lag, and upon the nature of the process of recovery.

Such a depressive effect is traditionally described as "emotional." It appears experimentally simply as a decrease in reflex strength, and may be defined as such. One characteristic is that it readily adapts out. As we have noted above, curves for extinction may be obtained that lack this cyclic effect. The essential condition is simply that time be provided for adaptation.

In Figure 6 we may inspect more easily the theoretical curves that have been fitted to the data. They are drawn as envelopes, upon the general assumption that the deviations are depressions. The above account of the fluctuation is based, of course, upon that assumption, since it requires the existence of a compensatory effect. Furthermore, the first short section is heavily weighted in each record, and the theoretical curves are permitted to lie considerably above the immediately following sections of the curves. This is also justified by the present interpretation of the fluctuation. Since we have assumed a time-lag, the first part of the curve must be free of the depressing effect. Moreover there is no reason to expect the first few recoveries from depression to be successful in reaching the extrapolated curve, since the generation of a depressing effect is in no way related to whether or not the rat is "on schedule." Eventually the theoretical curve is reached, because the depressive effect partially adapts out, and because the curve has meanwhile fallen off.

We can summarize the present analysis by listing the assumptions it involves:
1. The decrease in reflex strength observed during the process of extinction follows an approximately logarithmic course.

2. The essential condition for extinction (the interruption of the chain) leads also to a temporary ("emotional") change, which depresses the rate of elicitation.

3. The depressive effect shows a time-lag, which, in conjunction with a previously demonstrated compensatory increase in rate, produces a cyclic fluctuation.

4. The depressive effect is subject to adaptation.

These are all apparently reasonable assumptions, and they are the simplest set capable of accounting for the recorded curves. They are susceptible of independent investigation, and it should eventually be possible to give a complete quantitative description of the type of curve here reproduced.

Summary

A previously described method of investigating the formation of a conditioned reflex is suitable also for the study of experimental extinction. A procedure is described and four records are reproduced. The typical curve can be fitted with a logarithmic curve drawn as an envelope. The experimental curve, however, shows a characteristic fluctuation.

Several minor types of disturbance may be present in records of this sort. When due allowance has been made, the resulting inferred curve is wave-like in nature. The fluctuation is cyclic. It can be accounted for by assuming that the essential condition for extinction (the interruption of a sequence of reflexes leading, in this case, to the ingestion of food) generates a temporary "emotional" effect that depresses the rate of elicitation.

Curves are also reproduced showing extinction after reconditioning and extinction immediately after the original conditioning. In the latter case a very large ratio is observed between the number of times the reflex is elicited during extinction and the number of elicitations at which conditioning or over-conditioning could have been possible. It is also demonstrated that the increase in reflex strength observed in reconditioning is not simply a facilitative effect due to the consumption of the food obtained.

References


*Laboratory of General Physiology*
*Harvard University*
*Cambridge, Massachusetts*

**SUR LA VITESSE DE L'EXTINCTION D'UN REFLEX CONDITIONNEL**

(Résumé)

Une méthode déjà décrite pour l'investigation de la formation d'un réflexe conditionnel est aussi bonne pour l'étude de l'extinction expérimentale. On décrit un processus et reproduit quatre notations. La courbe typique peut être représentée par une courbe logarithmique dessinée comme enveloppe. La courbe expérimentale montre cependant une fluctuation caractéristique.

Plusieurs types moins importants de confusion peuvent être présents dans des notations de cette sorte. Quand on s'en est rendu compte, la courbe inférieure qui résulte ressemble aux ondes en nature. La fluctuation est cyclique. On peut l'expliquer en supposant que la condition essentielle de l'extinction (l'interruption d'une séquence de réflexes menant dans ce cas à l'ingestion de la nourriture) génère un effet "émotif" temporaire qui déprime la vitesse des réponses.

On reproduit aussi des courbes qui montrent l'extinction après un nouveau conditionnement et l'extinction immédiatement après le premier conditionnement. Dans le dernier cas on observe un très grand rapport entre le nombre de fois que le réflexe se présente pendant l'extinction et le nombre de réponses où le conditionnement ou trop de conditionnement auraient pu être possibles. On montre aussi que l'augmentation de la force du réflexe observée dans le nouveau conditionnement n'est pas simplement un effet facilitant dû à l'action de manger la nourriture obtenue.

**SKINNER**

**ÜBER DIE SCHNELLIGKEIT DES AUSSTERBENS (RATE OF EXTINCTION) EINES BE dingten REFLEXES**

(Referat)

Das schon beschriebene Verfahren zur Untersuchung der Gründung (formation) eines bedingten Reflexes lässt sich auch auf die Untersuchung der experimentellen Abschaffung anwenden. Es wird ein Verfahren beschrieben und es werden vier graphische Registrierungen (records) wiedergegeben. Der typischen Kurve lässt sich eine logarithmische Kurve anpassen, die als Umkreis (envelope) gezeichnet wird. Die experimentelle Kurve erweist aber ein karakteristische Schwankung (fluctuation).

Es können sich in Registrierungen dieser Art mehrere Arten von geringeren Störungen zeigen. Werden diese Störungen in Betracht
Rate of Extinction of Conditioned Reflex

gezogen, so ist die resultierende zu folgender (inferred) Kurve wellenartig. Die Schwankung ist zyklisch. Man kann sie durch die Annahme erklären, dass die zur Abschaffung wesentliche Bedingung (die Unterbrechung einer in diesem Falle zu der Einnahme von Nahrung führende Reflexenreihenfolge) (sequence of reflexes) eine affektive ("emotional") Einwirkung hat, die die Schnelligkeit der Hervorrufung hemmt.

Es werden auch Kurven wiedergegeben, die die Abschaffung nach Wiederbedingung (reconditioning) und die Abschaffung sofort nach der ursprünglichen Bedingung darstellen. Im letzteren Fall findet man eine sehr hohe Verhältniszahl (ratio) zur Äusserung der Beziehung zwischen der Zahl der Hervorrufungen (elicitations) des Reflexes während der Abschaffung und die Zahl der Hervorrufungen bei denen Bedingung oder Überbedingung (over-conditioning) hätten möglich sein können. Es wird ebenfalls erwiesen, dass die während der Wiederbedingung beobachtete Zunahme der Stärke des Reflexes nicht bloss ein fakultatives, durch die Einnahme der erhaltenen Nahrung verursachtes Ergebniss ist.

Skinner