Psychological Review
1971, Vol. 78, No. 1, 58–70

PUNISHMENT: METHOD AND THEORY [1]

PHILIP J. DUNHAM [2]
Dalhousie University

A methodological framework for the analysis of punishment is outlined. The methodology, which is called a multiple-response base-line procedure, serves two purposes. First, it raises a number of new questions about the properties of punishment. Second, it permits the examination of some untested assumptions found in traditional punishment theory. Initial evidence obtained with the multiple-response methodology questions the validity of traditional theoretical assumptions and suggests two simple rules for predicting the properties of various punishment operations.

When an aversive stimulus is contingent upon the occurrence of a particular response, a decrement in the probability of the response is usually observed. This procedure is typically called punishment and the decrement in response probability is called punishment suppression. The basic purpose of this paper is to delineate some fundamental problems with existing punishment theory and to suggest an alternative approach to the problem of punishment.

An Overview of Punishment Theory

Historically, there have been two fundamental assumptions used to explain the phenomenon of punishment suppression. The first of these assumptions was the strong version of the negative Law of Effect proposed by Thorndike (1913). Thorndike assumed that any painful or unpleasant event would weaken the response (or assumed S–R bond) which preceded that event. Thorndike (1932) subsequently rejected this notion and it has not enjoyed any serious attention since that time. The second fundamental assumption suggested to account for the punishment suppression phenomenon has been referred to as the *alternative-response assumption* (cf. Dunham, Mariner, & Adams, 1969). In its simplest form, the assumption states that the decrement in a punished response is caused by an increment in some alternative behavior.

All contemporary explanations of punishment suppression are specific elaborations of this alternative-response assumption. Those specific elaborations which have been most formalized fall into two major categories. These categories are referred to as single-process and two-process theories of punishment (cf. Solomon, 1964). The earmark of the single-process theory is the assumption that only one type of learning mechanism is involved in the development and maintenance of the alternative response during punishment training. Two types of single-process theory have been suggested and are differentiated in terms of suggesting either a classical or an instrumental conditioning mechanism. Estes and Skinner (1941), for example, suggested that emotional responses elicited by the punishing event are *classically* conditioned to stimuli which precede the punishing event. The classically conditioned behavior is assumed to compete with the punished response and cause the suppression. Miller and Dollard (1941) exemplify the instrumental conditioning version of single-process theory. They suggested that any response which is associated with the termination of the punishing stimulus will be *instrumentally*

[1] Research reported in this paper was supported by Project Grant APA-194 from the National Research Council of Canada. Many valuable suggestions have been made by different individuals during the preparation of this manuscript. The author is particularly grateful to N. J. Mackintosh, C. J. Brimer, and B. Moore for their critical comments.

[2] Requests for reprints should be sent to Philip J. Dunham, Department of Psychology, Dalhousie University, 1460 Oxford Street, Halifax, Nova Scotia, Canada.

conditioned as a response which escapes pain and competes directly with the punished response.

Two-process punishment theories specify two different learning mechanisms which are sequentially involved in the development and maintenance of the assumed alternative response. Dinsmoor (1954, 1955) and Mowrer (1947) are formal examples of the two-process explanation of punishment suppression. Dinsmoor suggested that the proprioceptive stimulus feedback from the punished response acquires secondary aversive properties via *classical* pairings with the primary aversive event. Any response which is *instrumental* in disrupting this chain of conditioned aversive stimulation will develop and be maintained as a response which competes with the punished response. Hence, two learning mechanisms, first classical then instrumental, are assumed to operate in the development of the alternative behavior. Mowrer's (1947) version of two-process theory substitutes the notion of conditioned fear for the notion of aversive stimulation conditioned in the Dinsmoor theory.

There are two basic implications of any version of the alternative response theory. First, there is the implication that some alternative behavior will develop and be maintained during punishment training. Second, there is the implication that this alternative response causes the reduction in the punished response. Presumably the former implication could be confirmed independent of the latter. But the latter could not be confirmed if the former were false.

In spite of the substantial amount of research on punishment in the last decade, there is no *direct* evidence to support either of the above implications of the alternative-response assumption. As Azrin and Holz (1966) have stated, the typical procedure in punishment research has been to infer the presence of the alternative response from the absence of the punished response. A minimal requirement for testing the assumption is the measurement of the alternative response independent of the phenomenon which it seeks to explain.

Two reasons can be suggested for the lack of direct evidence bearing on the alternative-response assumption. In the case of single-process theories, contingencies are specified which make the response elicited by the punishing event a prime candidate for conditioning during the suppression of the punished response. It is not a profound observation to note that punishment procedures which have traditionally been employed involve organisms, aversive stimuli, and apparatus which make it difficult to measure those responses which are suspected to participate in the relevant contingency. These "emotional" behaviors have not been recorded on impulse counters and this has prevented the accumulation of any evidence concerning changes in their probability during punishment training. The major problem with single-process versions of the alternative-response assumption at this point would appear to be the lack of an adequate methodology to test what would appear to be very testable implications.

With respect to the two-process theories, the problem is more serious. As Schuster and Rachlin (1968) have suggested:

> Because both the reinforcer and the response are unobserved and unobservable, the two factor theory of punishment poses a serious problem for the experimenter who wishes to test it: how can it be disproved? All the critical events are assumed to occur within the organism being punished [p. 784].

In addition to the specification of the critical contingencies "inside" the organism, it should be noted that *any* increase in some alternative behavior observed during punishment can be taken as support for the operation of the assumed internal contingency. Hence, the measurement of the unobserved behavior referred to by Schuster and Rachlin does not make the positions any more susceptible to disproof.

The picture which emerges from this brief overview of punishment theory and research is that there has been a lack of interaction between punishment theory and punishment data. The lack of any data relevant to the most fundamental theoretical assumptions has permitted the alternative response interpretations to persist as originally formulated.

In the next section of the discussion, a methodological approach to punishment will be described which can serve two functions in the context of the existing punishment literature. First, the methodology permits us to examine some questions about the properties of punishment which have not previously been considered. Second, it permits us to examine some previously untested implications of existing punishment theory.

A Methodological Approach to Punishment

The methodological approach to be suggested is called a multiple-response base-line procedure. It can be applied to a variety of problems in addition to punishment, and when viewed in the context of traditional methodology, it falls between a typical operant procedure, in which a single response is shaped under constraints, and typical ethological procedures, in which behavior is observed without external laboratory constraints. The most convenient way to describe the essential features of the methodology is to elaborate on a specific hypothetical example which is representative of several obvious variations on the basic approach. The hypothetical example should be noted with some care since it will be approximated in reality when experimental evidence is subsequently discussed.

Consider a small animal chamber with grid floor and three sources of enjoyment for the small rodent commonly called a Mongolian gerbil (*Meriones unquiculates*). The three items of interest in the chamber are a food bin with an unlimited supply of standard Noyes pellets, a drinking tube with unlimited supply of water, and adding machine paper which is threaded through a slot in the wall of the chamber.

The reader familiar with the behavior of this curious rodent will not be surprised to find that the gerbil will spend much of a half-hour daily session in the chamber shredding the adding machine paper, eating the food pellets, and drinking from the tube, in that order of preference. It is relatively easy to record the duration of each of these three behaviors during the session and convert these duration measures to a scale of response probability by dividing the amount of time spent in a particular response state by the total time possible (cf. Premack, 1965). Any type of behavior can be measured in terms of its duration, including the class of behavior which is labeled "doing nothing." With the appropriate manipulanda for a particular organism and control of the relevant parameters, a multiple-response steady-state base line of behavior, in which several measurable responses are observed to fill experimental time, can be established. No shaping or active manipulation of contingencies is assumed to be necessary. Assume, for the purposes of discussion, that our gerbil cooperated and filled the half hour with drinking ($p = .1$), eating ($p = .3$), and paper shredding ($p = .6$).

Once the multiple-response base line has stabilized, we are in the interesting position of actively manipulating the organism's environment and assessing the effects of such manipulations on *all* of the responses in the repertoire. Obviously, the manipulations of most interest in the present context are those which we call punishment operations; however, it is instructive to digress briefly and consider a simple manipulation like making a running wheel available. The most visible effect of making the running wheel available will be introduction of a running response into a repertoire of behavior which already fills experimental time. If running occurs, by definition, there will be a decrement in the observed probability of one or more of the existing responses in the repertoire. Curiously, there is little more than intuition to tell us how the organism will reorganize his response hierachy to accommodate the running response. Will he sacrifice a little bit of each of the existing behaviors? Will he select one response and sacrifice it for running privileges? If the latter, what is the rule of response selection?

If an aversive stimulus like electric shock is introduced, one is faced with roughly the same problem as that posed by the introduction of the running wheel. Shock, as an unconditioned stimulus, will define a certain probability of unconditioned behavior which must be assimilated into a response hierarchy. By definition, some decrement in the probability of one or more of the existing responses must take place. Again, we are not sure how the organism alters the existing preference structure to accomodate this additional behavior, but the rules which describe such alterations would be of importance in predicting the *immediate* effects of any variety of punishment operation on responses in the repertoire.

The point to be made with the two preceding examples is that some manipulations which can be used in the context of a multiple-response base line have the property of adding a response to the existing repertoire of behavior. Punishment procedures are one such manipulation, and it is suggested that procedures be arranged which permit the a priori specification of the response which will be added and subsequent measurement of that response during punishment training. In this respect, the selection of the gerbil was fortuitous. The response which we have observed to be associated with the introduction of shock is a vigorous biting and chewing of the grid floor. As reliable, if not as desirable, is an aggressive attack on another gerbil if the target animal is provided.

When one considers the variety of ways in which an aversive event like electric shock can be introduced into a multiple-response procedure, a number of empirical questions are generated which have not received previous experimental attention. A brief consideration of a few of these questions will illustrate the heuristic value of the multiple-response base-line procedure.

The traditional literature dealing with shock provides us with two major classes of operation which are called punishment and avoidance, with variations on each theme. In the context of the multiple-response base-line procedure, an alternative operational dichotomy is suggested. Specifically, the shock event can be delivered according to a program which makes reference to one or more responses in the repertoire, or according to a program without reference to the organism's response repertoire. The former includes traditional response-contingent punishment and Sidman avoidance operations; the latter includes various temporal schedules which are arranged independent of the organism's behavior. In future discussion, the terms response contingent and noncontingent will be used to describe these generic operational categories. Some measure of the comprehensiveness of any theoretical framework is provided by its ability to subsume data in both operational categories. Several of these operations, most often employed in single-response operant research, reveal interesting new dimensions when considered in the multiple-response context.

Consider the response-contingent punishment operation where one response is selected from the repertoire and the onset of that response is followed immediately by an instance of shock. The suppressive effects of this operation on the referent behavior are well known. However, the operation defines a version of Sidman avoidance contingency for all other responses in the organism's repertoire, including the response elicited by the shock event. In the gerbil example, response-contingent shock for eating also defines a very effective Sidman avoidance contingency for paper shredding. The more time spent paper shredding, the longer the interval between shocks and the fewer shocks received. Does the gerbil adjust his behavior in a manner suggested by the contingency?

Consider a fixed-interval noncontingent shock procedure. For a given density of shock, the probability of each response in the organism's repertoire will determine the number of shocks which are associated with that particular response. In the gerbil example, paper shredding occupies over half the experimental session, hence more shocks will arrive during paper shredding than during any other response in the repertoire. Of more interest, with shocks delivered at fixed intervals of time, that response which most consistently follows the shock event will most consistently predict the longest interval of "safety." Again, one must ask if the organism changes his behavior according to these implicit contingencies defined by the fixed-interval noncontingent schedule. A phenomenon reported initially by Morse, Mead, and Kelleher (1967) is perhaps relevant. These investigators reported that monkeys which are exposed to fixed-interval noncontingent shock will aggressively bite a rubber tube made available after each shock is delivered. The postshock elicited biting behavior would be the response in the organism's repertoire which most consistently predicts the longest interval of safety. As suggested by the preceding analysis, the animal changes its biting behavior accordingly. The amount of biting behavior increases during extended punishment training and eventually fills the shock–shock interval with a "scallop" in the rate of biting which appears at the end of the interval.

Consider also a Sidman avoidance procedure with a constant shock–shock interval and constant response–shock interval. Once we select a single response from the multiple-response base line as an avoidance response, all other responses in the repertoire are initially punished according to a noncontingent shock schedule, with the shortest possible intershock interval being the Sidman shock–shock interval. Which of the responses in the organism's repertoire should one select for the avoidance response for efficient learning? It would seem obvious that the response elicited by the shock would enjoy some advantages not enjoyed by the other responses in the repertoire. First, that response will predict the longest safe interval between shocks even if the avoidance contingency were not in effect. Second, it has a relatively high probability during initial training sessions, hence it will sample an implicit avoidance contingency quite often. Third, if it is explicitly given the avoidance property of delaying shock, there should be few responses which will compete with it for rapid learning. This suggestion is in line with Bolles' (1970) view that the organism's innate species-specific defense reactions (SSDRs) to an aversive stimulus are most readily learned as avoidance responses. I would suggest, however, that the rapid learning has nothing to do with their "innate defensive" properties. It is the property of immediately following the shock event which optimizes the conditions for avoidance learning—that is, not only does the response eventually reduce shock frequency, but it initially predicts safety for a longer period of time than any other response in the organism's repertoire.

The converse situation can also be suggested. If we give the elicited response the initial advantage of consistently predicting the longest safe interval of time, yet attempt to train a different response as an avoidance behavior, it would not be surprising to observe the animal attempting initially to develop the elicited behavior to the detriment of the avoidance response. This would be particularly true in the case of constant shock–shock intervals (e.g., the Morse, Mead, & Kelleher, 1967, phenomenon), as opposed to variable shock–shock intervals where the predictive properties of the elicited response are less evident. This also leads to the suggestion that a variable shock–shock interval will produce faster learning of avoidance behavior when the avoidance response is *other than the shock-elicited response* (cf. Bolles & Popp, 1964).

Finally, it is of some interest to consider the variations in temporal schedule which are possible with noncontingent shock. We have already discussed the fixed-interval case; now consider some implicit contingencies established by variable-interval cases. When programming shocks to occur at varying intervals in time, the distribution of intervals may be an important consideration. The typical variable-interval tape program is a rectangular distribution of intervals with some guaranteed minimum interval. It is typically described in terms of its arithmetic mean and delivers the shocks with a random sequence of intershock intervals. Of more interest is an exponential distribution of intervals which will deliver the same number of shocks at varying intervals, but is a continuous distribution. The probability of shock at any "moment" in time is a fixed value and the animal has no programmed interval which is guaranteed to be safe. In typical punishment studies where a single operant response is shaped, trained, and measured, the effects of shocks delivered according to fixed intervals, rectangular distributions of variable intervals, and exponential distributions of variable intervals, may be the same—the operant is suppressed. However, in the multiple-response procedure, very different effects may be found with the three distributions. As mentioned earlier, the elicited response enjoys the advantage of predicting a safe period most consistently in the fixed-interval shock schedule. This would be true to a lesser degree of variable-interval schedules with rectangular distributions. The elicited response would most consistently predict the minimum interval of safety programmed in that distribution. In the exponential distribution, however, no response consistently predicts a minimum safe interval. According to the program, shock is equally probable at any point in time. This removes those implicit contingencies for certain responses which are most evident in fixed-interval schedules. Again the question which must be asked is whether or not the organism adjusts his behavior according to the presence and absence of these implicit contingencies.

The preceding examples are intended to illustrate two points: first, the heuristic value of the multiple-response analysis in terms of generating testable questions; second, the sparsity of evidence relevant to these questions. This sparsity is understood when one recognizes the emphasis which has traditionally been placed on single-response measurement in both free operant and discrete-trial punishment research.

The discussion of the multiple-response methodology can be concluded with a brief consideration of the predictions which single-process and two-process punishment theories would make in the context of the hypothetical gerbil procedure. With respect to single-process theory, the use of any type of punishment operation will introduce the grid-biting response into the organism's response repertoire. This grid-biting response should be a prime candidate for both classical conditioning (Estes & Skinner, 1941) and instrumental escape conditioning (Miller & Dollard, 1941) since it follows the shock event more often than other responses. The predictions of the two-process versions of the alternative-response assumption must be considered separately in terms of predictions about the effects of contingent and noncontingent shock procedures. Basically, the two-process theories suggest that any response which disrupts the chain of discriminative stimulus conditions which precede the shock event will increase in probability and be maintained during punishment training. In the case of response-contingent punishment, these theories make the general prediction that any response other than the punished response might increase in probability. They do not provide us with a response-selection rule to tell us if one or all of the unpunished responses increases. In the case of noncontingent shock operations, the two-process theories fail to make any meaningful prediction. When a noncontingent shock operation is employed, every response in the organism's response repertoire will be part of a discriminative chain of stimuli which terminates with shock on some occasions. Hence, it is impossible to suggest that any response in the repertoire participates in an instrumental contingency which disrupts the chain of conditioned aversive stimulation.

Some Rules for Prediction

The predictions which the single-process and two-process explanations of punishment suppression make in the context of the multiple-response methodology have been discussed in the preceding section. Prior to considering some evidence, I would like to suggest some alternative rules for predicting the effects of a variety of punishment operations on the various behaviors measured in the multiple-response procedure. The rules are, at this point, tentative, and a systematic analysis of punishment in the context of a multiple-response base line may modify or contraindicate these initial suggestions.

Once the immediate change in performance caused by the introduction of an unconditioned behavior into the repertoire has taken place (shock is introduced), the organism has a hierarchy of responses described in terms of response probability. The two rules which attempt to predict the fate of each of the responses in the organism's repertoire during subsequent aversive training are the following: (*a*) That particular response in the organism's repertoire which is most frequently associated with shock onset and/or predicts the onset of shock within a shorter time than other responses will decrease in probability and remain below its operant base line. (*b*) That particular response in the organism's repertoire which is most frequently associated with the absence of shock onset and/or predicts the absence of shock onset for a longer period of time than other responses will increase in probability and remain above its operant base line.

Consider the spirit in which these two rules are formulated. I am suggesting that there are two basic contingencies of importance in any operation for delivering the shock. First, there is an instrumental punishment contingency. It has two important dimensions. A response can be more frequently associated with shock onset than other responses, or it can predict a given frequency of shock onset within a shorter period of time, or both. Thus, the two dimensions of the contingency are frequency and time. Second, there is an instrumental avoidance contingency. It also has two important dimensions. A response can be more frequently associated with the absence of shock onset, or can predict the absence of shock onset for a longer period of time, or both.

In traditional single-response operant methodology, the time not spent on the operant is usually assigned the label of "nonresponse" and is assumed to be a homogeneous mass of behaviors, (cf. Rachlin & Herrnstein, 1969). In spite of their tentative nature, the two rules described above should make it obvious that it may be of more value to recognize the "heterogeniety" of the nonreponse class —at least in terms of response probability differences. Further empirical analysis may reveal such factors as the sequential dependencies between different responses to be an important determinant of the organism's adjustment to aversive contingencies. This criticism of the single-response operant approach is equally applicable to the procedures which employ appetitive contingencies.

Some Evidence

Over the past year, the research completed by the author and his associates consists of a series of successive approximations to the multiple-response methodology which has been discussed throughout this paper. Many of the thoughts developed in the preceding discussion of punishment originated with a serendipitous observation made while studying the effects of punishment on key pecking in pigeons (Dunham et al., 1969). Dunham et al. observed that pigeons trained to peck a response key for grain on a variable-interval schedule missed the key and hit the wall area adjacent to the key with a certain number of pecks. The introduction of key-peck response-contingent punishment suppressed key pecking and increased the frequency of off-key pecks during punishment. These results suggested that an unpunished response, other than the shock-elicited response, will increase in probability and be maintained during punishment training.

Subsequently, a more deliberate approximation to the multiple-response methodology was attempted, taking advantage of the phenomenon of schedule-induced polydipsia (cf. Falk, 1966) in order to establish two steady-state responses which occupied a large portion of experimental time.[3] Falk (1966) reported that rats trained to lever press for food pellets on a variable-interval schedule will indulge in an excessive amount of drinking during an experimental session if a drinking tube is made freely available. Using Falk's standard procedure, two rats were trained one hour each day to the point of polydipsic drinking, and a mild .2-milliampere, .5-second shock was introduced which was contingent on the lever press response (Subject 1) or on the drinking response (Subject 2). With the relatively mild shock intensity to minimize the probability of the unconditioned (and unmeasured) responses elicited by shock, it was assumed that drinking would be the most probable unpunished alternative if lever pressing were punished, and that lever pressing would be the most probable unpunished alternative if drinking were punished. According to the rules outlined in earlier discussion, the punished response was predicted to be suppressed, and

[3] The author thanks Jon Little for his assistance with this experiment.

the most probable unpunished alternative in the response repertoire was predicted to increase in probability following some degree of initial disruption by the addition of shock-elicited behavior to the response repertoire.

The results of the initial polydipsia training phase, the first punishment training phase, a recovery phase, and a second punishment training phase are illustrated in Figure 1. During initial training, the animals developed polydipsic levels of water intake similar to those usually reported by Falk (1966). In the initial punishment phase, Subject 1 was shocked for lever pressing and Subject 2 was shocked for drinking. In both cases, the punished response was observed to be suppressed immediately and the unpunished alternative behavior was observed, following initial disruption, to increase to levels which exceeded the established prepunishment base-line level. The results obtained with Subject 1 were very surprising. In spite of the aberrant (polydipsic) base line of water intake prior to punishment, the response was observed to exceed that base line during punishment sessions. In the third phase of the procedure, the punishment contingencies were removed and recovery was observed for 14 sessions. With the removal of shock, all responses increased to levels even higher than preceding base lines. This overshoot following removal of punishment is typical of the recovery of punished responses (cf. Azrin & Holz, 1966). It is interesting to note that overshoot occurs with both the punished and unpunished alternatives, which were measured in the present procedure. Following the 14 days of recovery, punishment contingencies were again introduced—this time for the response which had *not* been punished during the first punishment phase. Basically the same phenomena observed during the first punishment phase were observed during the second. The punished responses suppressed and the unpunished alternatives were disrupted and then increased in probability. Unfortunately the increase in drinking observed in Subject 2 did not consistently exceed the prepunishment base line. This may be the result of the ceiling problem on water intake referred to earlier.

The final line of evidence from the author's laboratory work to date was an attempt to approximate the hypothetical gerbil procedure used as an example throughout this article. There were several reasons to extend experimentation beyond the case of two highly probable responses to three or more measured responses. First, the tentative rules specify that only one of the several unpublished alternative responses will increase in probability during punishment training. In order to test this prediction, more than one alternative must be measured. Second, a test of single-process theories is possible only if the unconditioned response introduced by the shock event is measured. Third, an alternative approach to the data in the polydipsia experiment would be the suggestion that the shock has some general arousal property which increases all unpublished responding. The demonstration of a single-response increase would make such an arousal mechanism less appealing than the type of contingency analysis suggested by the tentative rules.

To answer these questions, Kennedy Muyesu-Kaisha Munavi conducted a burdensome experiment in the author's laboratory which was a very close approximation to the hypothetical gerbil procedure described earlier. Nine gerbils were randomly assigned to one of three groups. Each gerbil was placed in a small response chamber with food bin, drinking tube, and adding machine paper for a daily half-hour session. The total duration of each of the three behaviors was recorded on elapsed-time meters by an observer looking at the animals through a one-way glass window. When a gerbil entered the chamber, it had been permitted access to only 80% of its normal ad lib intake of laboratory chow during the preceding 23½ hours, and the only water permitted was that available during the experimental session. Under these conditions, the gerbils managed to fill approximately 25 of the 30 minutes available each session with paper-shredding, eating, and drinking, in that order of preference. It should be noted that records were kept of the amount of food and water consumed during base-line observations. Subsequently, if eating

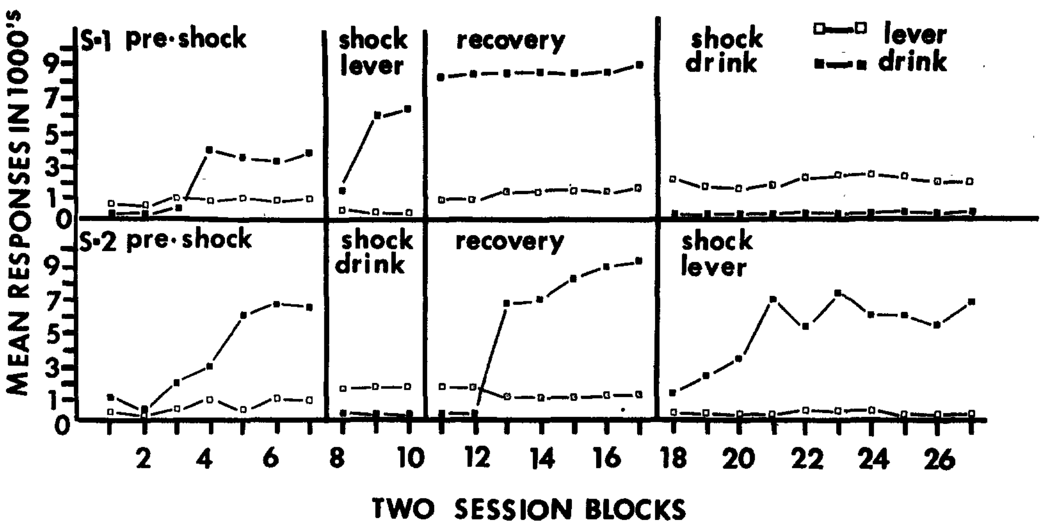

FIG. 1. Mean number of punished and unpunished responses plotted in two-session blocks for Subject 1 (S-1) and Subject 2 (S-2) in the polydipsia procedure.

was the punished behavior, the animal was permitted to make up any deficit in food intake during the experimental session in its daily ration one hour after the experimental session. Hence, there was no confounding of changes in either the total daily food intake or water intake with the introduction of the shock (cf. Dunham & Kilps, 1969).

After a base line of three behaviors was established, Group E was punished for eating, Group P for paper shredding, and Group D for drinking. The shock was a .2-milliampere, .5-second shock delivered at the onset of the referent behavior and continued at 2-second intervals until the referent response was stopped. At this point in the experimentation with gerbils, it was not known what elicited behavior or behaviors would be introduced with this shock intensity. After running four animals through the first session of punishment it was obvious that all animals were attacking the grid bars through which the shock was delivered. At this point a record was started of the duration of grid biting as a response which consistently occurred with the shock in all nine animals. Hence, we have the probability of grid biting during the first punishment session for only five of the nine subjects in this initial experiment. After the first session, grid biting was measured along with the three other responses during each session.

Prior to examining the results of the experiment, it is of some value to consider the predictions made by the rules when applied to this situation: (*a*) the rules suggest that the response most frequently and immediately associated with shock will decrease in probability and remain below its operant base line; it is the referent punished response in the response-contingent procedure; (*b*) there will be an immediate disruption of all responses in the repertoire immediately upon the introduction of the grid-biting response into the three-response repertoire which already nearly fills time; (*c*) after the initial disruption, the response in the repertoire which is most frequently associated with the absence of shock for the longest period of time will increase in probability to levels which exceed the prepunishment base line. This response should be that response which is most probable on the assumption that the most probable behavior will sample the implicit avoidance contingency most often.

Subsidiary predictions are also implicit in the rules. For example, if grid biting is not the most probable unpunished alternative behavior, it will drop out of the repertoire as the shock frequency declines (punished re-

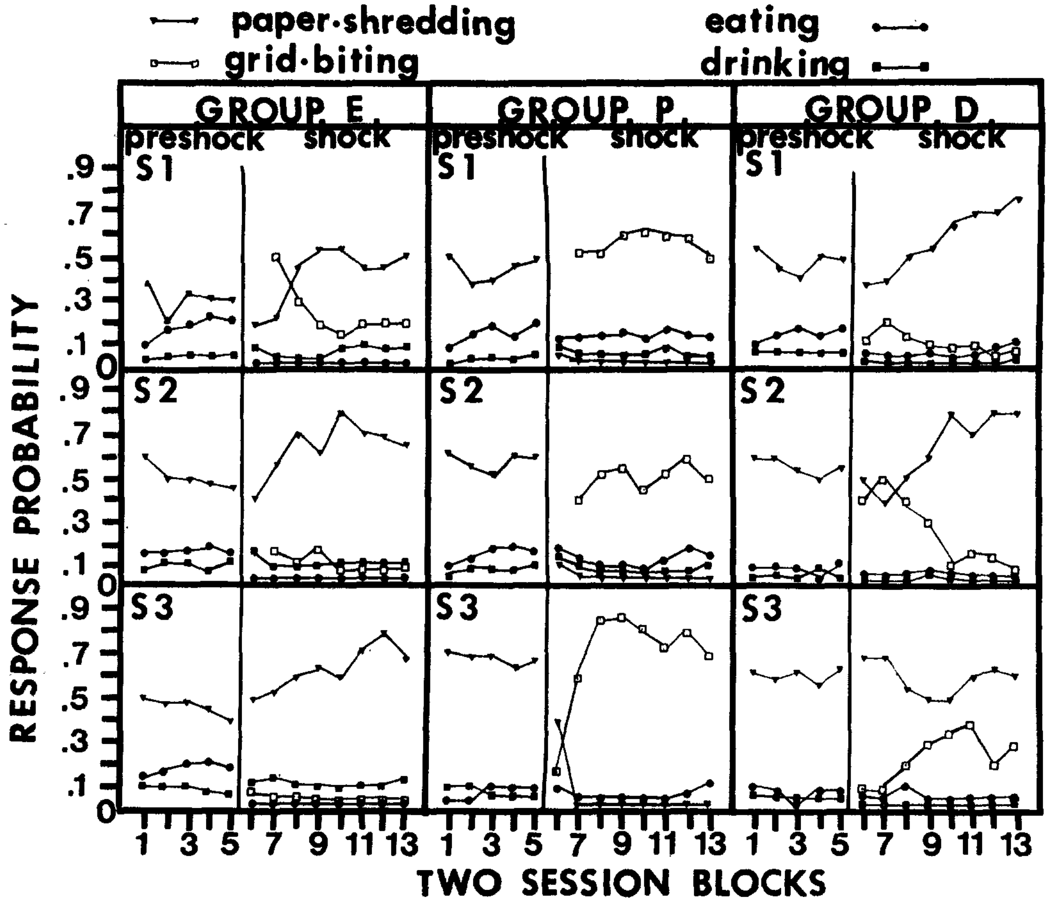

FIG. 2. Probability of punished and unpunished responses for the three subjects in each of the three groups in the gerbil procedure. (Note that all data points are two-session blocks with the exception of the first punishment session which is plotted as a single-session point.)

sponse suppressed). The latter suggestion is contrary to the predictions of the traditional single-process punishment theories. If, however, the grid biting response is the most probable of the unpunished alternatives, it will be maintained even in the absence of the shock UCS during punishment training.

The results of the gerbil experiment for each subject in each of the three groups are presented in Figure 2.

Consider first the two results which are common to all three groups and which are of primary interest in terms of the contingency rules. First, in all cases, the instrumental punishment contingency was sufficient to suppress the referent behavior. Second, and of more interest, in all cases, only *one* of the alternative responses in the repertoire was observed to increase in probability to levels which exceeded the baseline probability. These two results are taken as support for the operation of the contingencies outlined in preceding discussion.

Consider next the results specific to each group. Of particular interest in this context is the response which is selected to increase in probability during the punishment training. The subjects in Group E were punished for eating, and all three of the subjects were observed to increase the probability of paper shredding following an initial period of disruption. The subjects in Group P were punished for paper shredding, and in all three subjects the probability of grid biting increased during punishment training. Group D was punished for drinking and revealed inconsistent results. Subject 1 and Subject 2 increased the probability of paper shredding, while Subject 3 increased the probability of grid biting during punishment.

To what extent are the responses observed to increase in each group predicted by the contingency rules? The rules suggest that the response which most frequently samples the avoidance contingency in early sessions will be the most probable of the safe alternatives. In spite of the fact that the probability of the grid-biting response was not measured during the first session in four of the animals studied, it is safe to assume that two responses were the most probable unpunished behaviors during early punishment sessions. These are grid-biting and paper-shredding responses. Both would sample the avoidance contingency more frequently, and the grid-biting response which followed the shock event would predict the absence of shock for a longer period of time more consistently than the other responses which were not sequentially dependent on the shock event. For these reasons, we would expect all subjects to reveal an increase in either grid biting or paper shredding. However, there were some instances in which grid biting was initially more probable than paper shredding, yet paper shredding was observed to increase to levels above base line as grid biting dropped out (e.g., Group E, Subject 1). Further research, in which more precise measurements are obtained of the degree to which various responses sample both the frequency and time dimensions of the two contingencies, will provide a more critical test of the suggested rules and perhaps suggest refinements.

In addition to suggesting that the methodology and the contingency analysis are tractable approaches to the punishment problem, the gerbil data question the validity of the single-process versions of the alternative-response assumption. Stated quite simply, those cases in the present situation which revealed a gradual decline in the probability of grid biting over the course of punishment training are exactly the opposite to changes expected from traditional single-process assumptions. This is consistent with observations of shock-associated behaviors in other situations (cf. Dunham, et al., 1969). The emotional behaviors generally drop out within a very few sessions of response-contingent punishment training. Of course, this decrement in grid-biting behavior would not be expected in the case of noncontingent punishment in the sense that the UCS is maintained at a particular frequency independent of the changes in behavior.

The data also have implications for the two-process interpretations of punishment suppression. Basically, the very general prediction that some alternative behavior will increase is supported. It is questionable, however, to think that the increase observed in the alternative behaviors observed in the present experiments could account for the suppression in the punished response. The time course of the two transition processes are very different. The punished behavior is suppressed very rapidly relative to the initial disruption and slow rise of the alternative behavior.

The preceding evidence involves the response-contingent punishment operation. The author has not yet started an empirical analysis of the noncontingent punishment procedures. There are, however, several studies in the literature which are relevant to the predictions made by the rules discussed in this article as applied to noncontingent procedures. As indicated in earlier discussion, the organism cannot, by definition, reduce the shock frequency when a noncontingent operation is employed. This leaves the temporal dimension of the contingency in an important role. The response which consistently follows the shock event will predict the absence of the next shock onset for a longer length of time than any other response in the repertoire. The phenomenon reported by Morse et al. (1967) in which an elicited behavior develops and modulates under fixed-interval noncontingent shock suggests that the organism is sensitive to the temporal dimension of the avoidance contingency. More recent evidence reported by Powell and Creer (1969) can be subjected to a similar interpretation. These investigators delivered noncontingent shock at the rate of 100 shocks per session to pairs of rats and recorded the frequency of aggressive attacks per session. The authors did not specify the intershock interval or the session length in their procedure. Ten successive days of noncontingent shock sessions were conducted and measures of aggressive behavior indicated that an increase in the amount of aggression occurred over the course of the experiment. Although the experiment was designed to determine the effects of maturation (among other things) on aggressive behavior, it is suggested that the 10 days of maturation were confounded with 10 days of avoidance training during the experiment. Specifically, the aggressive behavior of these rats predicted the absence of shock onset for a longer period of time than any of the other (unmeasured) behaviors in the organism's repertoire. Similar to the monkeys in the Morse et al. procedure, the rats in the Powell and Creer experiment are assumed to have recognized this avoidance contingency and adjusted the behavior appropriately.

Assuming for the moment that the avoidance interpretation of the Powell and Creer experiment is correct, the present rules have implications for the development of aggression; namely, that certain stimuli will elicit aggressive behaviors, and under a variety of procedures used to deliver the stimuli, the aggressive behavior will develop as an avoidance response and be maintained in the absence of the primary UCS. Under other conditions, the avoidance contingency is not present and one should observe aggression at the unconditioned level. I know of no theoretical account of aggression which implicates an avoidance contingency in the development and maintenance of aggression in animals. Additional work with temporal schedules of aversive stimulation which eliminate this avoidance contingency should reveal the extent to which it is involved in the conditioning of the aggressive behavior of various species.

Some Concluding Remarks

In the absence of more data, it would be premature to attempt to elaborate on the type of theoretical mechanism which is implied by the empirical rules which have been outlined in the preceding discussion. I would like, however, to conclude by suggesting briefly the general directions which one might take in the development of such a mechanism. First, the data and arguments which have been discussed suggest very strongly that there is little to recommend those traditional theoretical accounts of punishment which start from an alternative response assumption. In place of this assumption, the rules which have been outlined imply a symmetrical conditioning mechanism in which (*a*) those responses which predict the aversive event are actively inhibited and (*b*) those responses which predict the absence of the aversive event are actively excited.

History provides at least two examples of symmetrical excitatory and inhibitory mechanisms which deal with appetitive and aversive events. If one is biased toward Pavlovian conditioning mechanisms, it is possible to conceptualize the multiple-response methodology as a *multiple-CS* methodology in which different responses are viewed as different CSs which predict, to varying degrees, the presence or the absence of the unconditioned stimulus (aversive or appetitive). Once conceptualized in terms of Pavlovian operations, some version of Pavlov's (1927) concepts of inhibition and excitation can be developed to explain the effects of punishment on performance. Konorski's (1967) inhibitory and excitatory processes and some recent modifications of these ideas (e.g., Maier, Seligman, & Solomon, 1969) should, for example, be considered in the context of the multiple-response punishment procedure.[4]

[4] A basic assumption in Konorski's theorizing is the notion that noxious stimuli such as shock elicit a drive state, fear, which has general inhibitory effects on "preservative" drive states such as hunger, thirst, etc. Punishment would be predicted to have suppressive effect on *all* responses motivated by "preservative" drive states. In the experiments conducted thus far, the inhibitory effects which we have observed are specific to the punished response. For example, when the gerbils were punished for eating they continued to drink the usual amount each session. Hence, if Konorski's inhibition mechanism were to be considered in the context of punishment, the inhibitory properties

Alternatively, if one is biased toward instrumental conditioning mechanisms, it is possible to conceptualize the multiple-response methodology in terms of instrumental contingencies in which each response has either rewarding or punishing consequences. Once conceptualized in terms of instrumental operations, some version of Thorndike's (1913) positive and negative Law of Effect can be elaborated upon as a symmetrical reinforcement mechanism. Rachlin and Herrnstein (1969) have recently made such a suggestion.

In either case, it is hoped that the multiple-response base-line procedure will help to restore the interaction between punishment theory and punishment data which has been lacking in the punishment literature.

---

of shock would have to be more "drive" or "response" specific than he has assumed. Estes (1969, p. 69), in his more recent theorizing, makes an assumption very similar to Konorski's. Specifically, he suggests that the activation of negative drive systems (e.g., shock-produced attack or flight) results in the inhibition of positive drive systems, which, in turn, accounts for the decrement in punished responding. Again, it should be noted that the gerbil data suggest that the only positive drive system which appears to be permanently inhibited during punishment is the specifically punished drive system—or its associated response.

REFERENCES

AZRIN, N. H., & HOLZ, W. C. Punishment. In W. K. Honig (Ed.), *Operant behavior: Areas of research and applications.* New York: Appleton-Century-Crofts, 1966.

BOLLES, R. C. Species-specific defense reactions in avoidance learning. *Psychological Review,* 1970, **77,** 32–48.

BOLLES, R. C., & POPP, R. J., JR. Parameters affecting acquisition of Sidman avoidance. *Journal of the Experimental Analysis of Behavior,* 1964, **7,** 315–321.

DINSMOOR, J. A. Punishment: I. The avoidance hypothesis. *Psychological Review,* 1954, **61,** 34–46.

DINSMOOR, J. A. Punishment: II. An interpretation of empirical findings. *Psychological Review,* 1955, **62,** 96–105.

DUNHAM, P. J., & KILPS, B. Shifts in magnitude of reinforcement: Confounded factors or contrast effects. *Journal of Experimental Psychology,* 1969, **79,** 373–374.

DUNHAM, P. J., MARINER, A., & ADAMS, H. Enhancement of off-key pecking by on-key punishment. *Journal of the Experimental Analysis of Behavior,* 1969, **1,** 156–166.

ESTES, W. K. Outline of a theory of punishment. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior.* New York: Appleton-Century-Crofts, 1969.

ESTES, W. K., & SKINNER, B. F. Some quantitative properties of anxiety. *Journal of Experimental Psychology,* 1941, **29,** 390–400.

FALK, J. L. The motivational properties of schedule induced polydipsia. *Journal of the Experimental Analysis of Behavior,* 1966, **9,** 19–25.

KONORSKI, J. *Integrative activity of the brain. An interdisciplinary approach.* Chicago: University of Chicago Press, 1967.

MAIER, S. F., SELIGMAN, M. E. P., & SOLOMON, R. L. Pavlovian fear conditioning and learned helplessness. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior.* New York: Appleton-Century-Crofts, 1969.

MILLER, N. E., & DOLLARD, J. *Social learning and imitation.* New Haven: Yale University Press, 1941.

MORSE, W. H., MEAD, R. N., & KELLEHER, R. T. Modulation of elicited behavior by a fixed-interval schedule of electric shock presentation. *Science,* 1967, **157,** 215–217.

MOWRER, O. H. On the dual nature of learning a reinterpretation of "conditioning" and "problem-solving." *Harvard Educational Review,* 1947, **17,** 102–148.

PAVLOV, I. V. *Conditioned reflexes.* (Trans. by G. V. Anrep) New York: Dover, 1927.

POWELL, D. A., & CREER, T. L. Interaction of developmental and environmental variables in shock-elicited aggression. *Journal of Comparative and Physiological Psychology,* 1969, **69,** 219–226.

PREMACK, D. Reinforcement theory. *Nebraska Symposium on Motivation,* 1965, **13,** 123–188.

RACHLIN, H., & HERRNSTEIN, R. J. Hedonism revisited: On the negative law of effect. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior.* New York: Appleton-Century-Crofts, 1969.

SCHUSTER, R., & RACHLIN, H. Indifference between punishment and free shock: Evidence for the negative law of effect. *Journal of the Experimental Analysis of Behavior,* 1968, **11,** 777–786.

SOLOMON, R. L. Punishment. *American Psychologist,* 1964, **19,** 239–253.

THORNDIKE, E. L. *Educational psychology.* Vol. 2. *The psychology of learning.* New York: Columbia University, Teacher's College, Bureau of Publications, 1913.

THORNDIKE, E. L. *The fundamentals of learning.* New York: Columbia University, Teacher's College, Bureau of Publications, 1932.

(Received April 21, 1970)