

Reinforcement Theory¹

DAVID PREMACK²

University of Missouri

Along with imposing various constraints upon the subject in order to study reinforcement, we have been conducting other studies in which the subject is as little constrained as possible. What the rat does when unconstrained may be looked upon as the operating characteristics of the species; in time, the principles of reinforcement may prove to be illuminated by, if not derivable from, a proper understanding of the operating characteristics. Accordingly, we will start with a brief account of the unconstrained rat and then turn to the animal that is constrained in the various ways which a study of reinforcement demands.

FREE RESPONDING AS A BASE STATE: OPERATING CHARACTERISTICS

To obtain a comprehensive account of the operating characteristics, we perform what might be called an anti-experiment: vary nothing and see what the animal varies. That is, place the organism in as undemanding and invariant an environment as possible and observe the manner in which the animal varies despite the invariance of the environment. Environments come close to leaving the animal alone when they drop three standard constraints: contingencies, deprivation, and gradients. If there are no gradients of temperature, light, or noise, any activity cycles found can be attributed to the animal. Similarly, if stimuli such as food and water are continuously available, rather than

¹ The research reported here was aided by grants from the National Science Foundation, the National Institute of Mental Health, and the University of Missouri Graduate Research Council. I am indebted to Walter Kintsch for criticism and helpful comments.

² Present address: University of California, Santa Barbara.

being constrained, on the one hand, by a deprivation schedule or, on the other, by an instrumental response requirement, then any variation in the animal's responding to these stimuli can be ascribed to the animal.

Four rats were maintained for 150 days in an environment approximating the one described.³ The animals were housed individually in standard-size cages, each of which provided an activity wheel, a food source, and a drinkometer. Continual illumination was used, along with the usual controls for ambient sound and temperature. Throughout the 150 days of observation, the animals had continuous access to food, water, and wheel, except for brief periods of daily maintenance.

The results may be best summarized by starting with the smaller units of behavior and working toward the larger. A fair amount has been written about the invariance of the lick rate: when rats lick, they lick about 7 times/sec. (Stellar & Hill, 1952). We have found 18-day-old rats, licking for the first time, to lick 7 times/sec. (Schaeffer & Premack, 1961). But it is less widely considered that an apparently comparable invariance is found in other behaviors; indeed, it would be surprising if the invariance were confined to licking. Other behaviors do not offer as convenient a topographic unit as the lick, but it is not strictly necessary to measure recurrence anatomically in order to establish that certain movements of the animal may occur at constant rates. For example, rats appear to eat standard 45-mg. Noyes pellets in about 6.5 sec. As in the case of the lick, this appears to be independent of deprivation: ad lib fed and starved rats both consume the pellet in about 6.5 sec. Similar constancies appear to obtain for running. Skinner (1933) observed some years ago that in a properly balanced wheel, the rat's paws contacted the running surface a constant number of times/sec. The intervening years have not dimmed these constancies; rats complete most 360-degree turns in our wheels in about 1 sec. (Premack & Schaeffer, 1962). If locomotion differs from eat and drink, it is in occurring in more than one form. But each form appears to be nearly invariant: locomotion may thus be a family of con-

³ R. W. Schaeffer collected these data, assisted by G. Collier and myself.

stant-speed gaits. Grooming, too, appears to consist of a small set of stereotyped motor sequences, each with a fixed rate of occurrence. Finally, the copulatory movements of the male rat may be subject to a comparable stereotypy. Although only a few of these cases are based on measurement, and others—grooming and copulation—on impression alone, it is not too great an inductive leap to wonder if, in the rat, all recurrent behaviors do not occur at constant momentary rates.

Dethier (1964) has aptly dubbed an apparently similar state of affairs in certain insects the "little motor theory of motivation." We may view the fetal rat as moving down an assembly line, receiving first a motor from one bin, a little later on a motor from another bin, the bins marked "eat," "drink," "run," etc. Each motor has only one operating speed, and this speed is the same for all motors in the same bin. That is, at the level of the "little movements," there are not even individual differences. Seven licks per second is not an average figure, but a constant from rat to rat.

Individual differences are found up a level, not in the speed at which the motor operates, but in the durations for which the motors operate once they have started. The story appears to be much the same for eat, drink, and run, and so may be told for all of them in terms of any one of them. For example, the rat eats (or drinks or runs) throughout the day in a series of bursts—periods in which the behavior is essentially unbroken.⁴ Dura-

⁴ Bursts were originally defined as the ingestion of a 45-mg. pellet, a tongue lap, or 90 degrees of wheel revolution separated from any other like instance by at least 8 sec. To examine the possibility that definition of burst might affect the outcome of subsequent correlational analyses, certain of the analyses were repeated using different burst definitions. Bursts were redefined as like instances separated by as little as 0.1 sec. and by as much as 32 sec.; for this appreciable range, burst definition had no essential effect upon the conclusions reported here. A condition which may have more effect upon the present conclusions is that the data be for animals that are thoroughly adapted (and perhaps kept in constant light rather than constant dark). We find rats to drop their activity cycles after several weeks of such maintenance, and thenceforth to distribute their eat, drink, and run more or less randomly throughout the day. This may or may not be essential for the conclusions reported here; in any case, our animals tended to be stationary. As regards specifically eating, more recent data based upon bulk food obtained at the end of a "tunnel" confirm the present data based upon pellets obtained by a bar press.

tions of the individual bursts are definitely not constant, but vary over a considerable range. However, the average of the distribution of burst durations for any behavior is reasonably stable for the individual rat from day to day. Also, the three behaviors tend to order on the basis of the average burst duration, eat being the longest and run the shortest. Thus, for any one behavior it is possible to tell one rat from another by the average duration of the burst—how long the “little motor” operates once it has started—and for any one rat tell one behavior from another on the same basis. Some indication of the range of the distributions and stability of their means and variances is provided in Table 1.

TABLE 1
MEANS AND VARIANCES OF BURST LENGTH FOR EACH BEHAVIOR
PER SIX-HOUR PERIOD

		Period I	Period II	Period III	Period IV
<i>Drink</i>					
S1	Mean	4.4	4.3	5.4	5.0
	Variance	7.5	7.3	9.6	6.2
S2	Mean	10.8	8.0	12.8	10.2
	Variance	36.4	38.8	26.2	37.1
S3	Mean	7.0	5.7	7.7	4.3
	Variance	18.9	4.0	16.1	21.1
S4	Mean	5.9	6.5	9.0	8.4
	Variance	12.4	19.3	18.6	21.0
<i>Eat</i>					
S1	Mean	12.0	11.1	10.0	9.2
	Variance	87.7	76.3	61.4	68.6
S2	Mean	10.6	10.7	12.9	10.7
	Variance	130.0	106.0	119.3	88.1
S3	Mean	21.0	15.0	20.7	16.9
	Variance	140.3	146.2	205.1	137.0
S4	Mean	4.6	3.7	4.2	3.9
	Variance	27.4	19.3	22.5	25.4
<i>Run</i>					
S1	Mean	4.3	4.0	4.2	3.7
	Variance	18.7	13.9	15.9	15.5
S2	Mean	3.4	4.1	3.6	4.2
	Variance	7.2	23.8	7.0	9.0
S3	Mean	3.7	3.8	4.8	4.8
	Variance	19.7	14.4	25.9	23.5
S4	Mean	3.8	4.1	4.8	3.5
	Variance	25.4	23.7	30.9	15.2

Is it possible to account for the variation in duration of the individual bursts? Since the individual bursts vary widely in duration, it would seem reasonable that factors could be found which would account for this variation. The rat sometimes engages in a burst of eating that lasts for five or six minutes, while on other occasions it may eat a "meal" that lasts only a matter of seconds. How long it waits between bursts also varies considerably. It would seem reasonable to suppose that if a current "meal" were greater than average, the rat would wait longer than average before eating again and/or would eat a "meal" smaller than average.

Again the answer is very similar for all three behaviors and thus may be told in terms of any one of them. All seemingly reasonable suppositions as to what may control burst duration turn out to be unfounded; at least we have been unable to collect evidence in support of them. For example, auto-correlations on adjacent pairs of like bursts are typically of the order of about .04. This is borne out by the Von Neuman ratio, which is 2 for independent events and for which we obtain 1.96, 1.97, and 1.95 for the relation between durations of successive pairs of eat, drink, and run, respectively (Kintsch & Premack, 1965). Likewise, the correlation between duration of the individual burst and duration of the preceding nonresponding interval of like kind—for example, "meal" and preceding noneating interval—is of the order of about .20. The only correlations we have found of any magnitude are those between duration of the burst and total duration of like responding in the preceding 15 min. However, these largest of all correlations leave approximately 87 to 98 per cent of the variance in burst duration unaccounted for.

There would thus appear to be an appreciable random component in the duration of the individual bursts. This conclusion is so far based on negative evidence; it will remain so until we are able to go beyond suggesting a random process and can specify what kind of random process. (It is exactly at this point that the need for massive data is felt.) In the meantime, there remains the fact that, given that the animal has started to

respond—the motor has begun to operate—we are unable to find any factors that provide appreciable help in predicting when it will stop. This state of affairs would appear to be incompatible with standard reactive inhibition-type models, for rest and amount of responding are not playing their assumed roles. One condition which this state of affairs does support is the propriety of using response duration as an estimate of response probability. The relevance of this last implication will become evident when we turn to reinforcement, and we will leave its development for that time.

Although durations of the individual bursts are unpredictable, the average of the distribution of burst durations can be manipulated in the standard deterministic fashion. For example, by increasing the sucrose content of the food, we can shorten the average duration of bursts of eating, but without affecting the unpredictability of the duration of the individual bursts. Doubtless other parameters can be found that will affect the average of the process without simultaneously affecting predictability of individual events. Probably what is most surprising here, by (risky) analogy to physics, is that the unpredictability is for macroscopic events—the behavior of the whole organism. A comparable unpredictability for neural events would be more easily accepted.

What behavior state the rat enters into is in marked contrast to how long it stays there, for the data reveal strong dependencies between current burst and kind of immediately preceding burst. Thus, whether the rat runs, eats, or drinks depends upon what it did previously. Additional though weaker dependencies are found for two-preceding and three-preceding bursts. What motor operates thus depends upon which ones recently operated. Kind of burst will not predict duration, however, any more than duration will predict duration. Thus, while a drink is more likely to follow a run than would occur were the two classes independent, there is nonetheless no correlation between duration of the preceding run and duration of the drink that follows. Furthermore, the number of drinks following long runs is about the same as the number that follow short runs. Thus, kind of burst

will predict only kind, while duration will predict neither duration nor kind.

To summarize *tentatively* the operating characteristics of the rat:

1. The motor sequences that comprise the recurrent behaviors occur at fixed rates.

2. Duration of the individual bursts, that is, when an animal will stop a given behavior, is to a large extent unpredictable.

3. The behavior the rat enters is dependent upon the behaviors it previously entered.

4. Much of the unconstrained increase in any behavior is produced by shortening the pauses between successive bursts, that is, by entering the state more often (Premack & Schaeffer, 1962, 1963).

These are not offered as an account of the rat's response system, but as tentative descriptions representing the present state of our knowledge. We are now in a position to turn to reinforcement, having at least some idea of the natural response tendencies of the animal before we operate upon them with the constraining procedures of reinforcement.

REINFORCEMENT: INTRODUCTION

The traditional account of reinforcement provides a nice introduction to the present one, for the latter is a simple transform of the former in the following sense. First, the basic assumptions of the traditional account appear to be empirical propositions with a confirmation status so high that further test is unnecessary; actually, the assumptions were little more than common sense, which is to say, they were never tested. Second, when tested, they proved to be invalid. Thus, all of the major assumptions of the traditional account have essentially the same logical character. They appear to be propositions involving empirical claims, claims so weak as to be essentially trivial. It is only when the propositions are seen to be nonempirical that they become nontrivial.

We may start with the standard definition according to which a reinforcer is any stimulus that, given a certain relation to a response, produces a change in the frequency of that response.

Since presumably it is only philosophers who delight in null classes, tradition hastened to add that, of course, there are such stimuli. From there it seemed only reasonable to add: there are other stimuli that do not have this property. The combination, there are some stimuli that have this property and there are others that do not, will be recognized as a familiar version of the first of the several assumptions that make up the traditional descriptive account (e.g., Skinner, 1938, p. 62).

The second empirical claim is that reinforcers are trans-situational (Meehl, 1950). Inevitably, the discovery that any stimulus is effective will be made with some particular base response; the claim of trans-situationality is that a stimulus effective with one response will be effective with all responses (and thus that the discovery could have been equally well made with any response). The classical example is food; it is said to reinforce all responses.

A third empirical claim is that there are two classes of responses, one that is reinforcing but not reinforceable, another that is reinforceable but not reinforcing. This is less an explicit assertion, though versions of it can be extracted from traditional discussions of goals and drives, than a conclusion that can be reached by examining a sufficient body of experiments. For example, eating and drinking were for many years used only in the role of reinforcer, never as the events that were reinforced; conversely, bar pressing and running were used only as "instrumental activities," never as reinforcers. The segregation here was impressive, there being no instances of crossover between these would-be categories in about forty years of experimentation.

It is important to recognize that the above three assumptions were not viewed as a *theory* of reinforcement, but rather as a first-order, low-level description of the field. Theorists differed as to whether or not reinforcement was coterminous with, say, drive reduction, or whether it was necessary for learning, but they did not differ in their acceptance of the three assumptions.

Are these assumptions, which certainly give the impression of being empirical, actually members of that class in the standard sense of being abstractions from test evidence?

In order for the assertion concerning two classes of stimuli

to be an empirical one, the literature must show some body of experiments in which, for any species, some stimuli were shown to change the frequency of a response while other stimuli, given the same temporal relation to the response, failed to produce a change. This kind of functional division of the environment is clearly implied by the classical assumptions, but it is equally clear that no actual division ever took place. That is, it is not possible to list for any species any set of events that have been shown not to be reinforcing. Nor is this omission the result of being hamstrung by the null hypothesis. The tests were not made because the claim of two classes was never seen to be a claim of any empirical consequence. The fact was too trivial to merit test.

The assumption of trans-situationality cannot be met with the same charge of being wholly nonempirical, for food has been shown to reinforce a number of responses. Food has at one time or another reinforced the bar press, running down an alley, the string pull, and pushing a marble into a hole; doubtless others could be added. But the necessity of piecing together these outcomes from different experiments, which normally would be disallowed, shows that the comparisons are post hoc; the issue was so little seen as empirical that probably not even a master's thesis was devoted to it.

The third assumption could not have been tested (except by accident), for we have already observed that it was tacit, failing even to attain to the status of a necessary condition for being tested, viz., to be made explicit. With a few quite recent exceptions, no attempts were made to reinforce eating or, on the other side of the coin, attempts made to reinforce with the bar press. These events were kept in their proper places. It is no surprise, therefore, that when Meehl's Martian arrived in the animal lab, primed to execute an epistemological reconstruction of reinforcement, he somehow managed to get the responses "right," i.e., managed to reinforce running with eating rather than, as a Martian might, the reverse. (Was the Martian earthling? He seemed a little too knowing.)

The core assumptions of the traditional account wear empiri-

cal masks which, when they are removed, reveal fairly commonsensical assumptions. The assumptions were not tested; it had already been decided what kinds of events were reinforcing and what kind not. Hence the surprise when lights, sounds, puzzles, etc., proved to be members of the class.

Two choices were given by the "new" reinforcers, either to enlarge the drive category, admit some unforeseen events, and essentially retain the traditional logic of reinforcement, or to drop the standard approach and seek another one.

One reaction to the sudden plethora of reinforcers—lights, sounds, puzzles where a moment ago there had been only food—was to pronounce everything effective, to declare all stimuli reinforcing. But this blanket denial of the former categorical assumption is not one for which any appreciable evidence was ever collected.

The approach we have taken is to observe that reinforcement involves a *relation*, typically between two responses, one that is being reinforced and another that is responsible for the reinforcement. This leads to the following generalization: of any two responses, the more probable response will reinforce the less probable one. Given the generalization, the view of which it is the core can now be set forth in an orderly manner:

1. Anatomically different responses can be compared directly.
2. For any pair of responses, the more probable one will reinforce the less probable one.
3. An indifference principle holds such that the reinforcement value is determined by response probability independent of parameters used to produce the probability or kind of response that manifests the probability.
4. Reinforcement is a relative property. The most probable response of a set of responses will reinforce all members of the set; the least probable will reinforce no member of the set. However, responses of intermediate probability will reinforce those less probable than themselves but not those more probable than themselves. Intermediate members of the set thus both are and are not reinforcers, depending upon the relative probability of the base response.
5. The reinforcement relation is reversible. If the probability of

occurrence of two responses can be reversed in order, so can the reinforcement relation between the two responses.

The logical status of these five assumptions differs from one assumption to the next, and they are presented as a group in the interests of communication rather than for their logical homogeneity. For example, 2 is the central empirical claim, 1 is presupposed by 2, while 4 and 5 are revealing implications of 2.

For the generalization to be tenable, no more than that, it must be possible to compare anatomically different responses. Obviously, it is not possible to talk about the effect that a more probable response may have upon a less probable one if it is not first possible to establish the order of their probabilities. Is it possible to compare anatomically different responses? Tradition may give the impression that it is not, but if so, the impression is misleading, for tradition has not decided against the comparison so much as it has simply never involved itself in issues that required such comparisons.

RESPONSE METHODOLOGY

Standard response methodology deals only with what may be called *within-response* comparisons. One and the same response class is compared under different conditions. Approximately 99 per cent of psychological data would appear to be of this sort. As a simple example, consider the classical question of whether rats run more when hungry than when not. Along with more complex issues, this simple question is answered by a within-response comparison. With the use of an activity wheel, a run may be defined as, for example, a 360-degree turn, provided only that a run is counted by this same unit both when the rat is and is not hungry. The use of 360 degrees is, of course, quite arbitrary—359 degrees might be used as well, or even 35 degrees—but the arbitrariness is not injurious to the comparison, for the only criterion in the case of within-response comparison is consistency. That is, whatever the response unit used, it must be the same for the several conditions of the comparison. The criterion for *between-response* comparison is, as we shall see, decidedly

stronger, although this is no criticism of standard response procedure, for there is no advantage in making criteria stronger than the performance of a task requires.

A further reason for possibly supposing that comparisons between different responses may not be possible is the sheer number of different measures that are used, along with the evident fact of their incommensurability. For example, volume, number of pellets, licks, degrees of turn, number of bar presses, and speed of running are all in use, and quite legitimately, although they are plainly incommensurable. Which of them are to be used for between-response comparison? They cannot all be equally suitable, and yet we appear to have no rule for deciding among them.

In fact, a rule is easily found, for while the criterion for within-response comparison is merely consistent application of *any* unit, the criterion for between-response comparison is that *all* responses be measured by the *same* unit. This requirement is not special to responses, of course, but is simply the general requirement for the comparison of any set of items. Thus apples and eggs are commensurable in the case of weight because the same measurement procedure can be applied to both. Similarly, diverse responses will be commensurable, provided they can be measured by the same unit.

A little reflection will show that time, response duration, will fulfill the requirement nicely. Given an operational definition of the onset and offset of a behavior, a clock sensitive to the fulfillment of the operational definition will determine the duration for which the organism is in a state of the designated responding. The clock will not tick more rapidly because the rat is copulating rather than merely eating; the unit of measurement will be invariant from one behavior to the next. Moreover, time is extrinsic to all behaviors and thus can be applied universally, in contrast to all anatomical units, such as the lick, which are intrinsic to given motor sequences and can be used to measure only those sequences composed of the unit in question.

Time not only fulfills the requirement of measurement by a common unit, thus permitting comparison of anatomically diverse

responses, but it takes priority over traditional response criteria such as common-effect-upon-the-environment or anatomical similarity. In fact, these traditional criteria do not provide sufficient conditions for commensurability, and the attempt so to use them can lead to malpractices. Consider two response classes whose members are either substantially alike anatomically or which have like effects upon the environment but have different average durations. Can the probabilities of the two classes be estimated comparatively from their frequencies? In which case is the organism more apt to be in a state of responding, when it makes 90 1-sec. responses or 10 15-sec. responses? This example alone should make it clear that having a comparable effect upon the environment, such as getting a bar down, is not a sufficient basis for comparing the frequency of two events when the frequency is to be used as an estimate of probability. Comparable average durations of the events in question is at least a necessary condition, and we would argue that it is a sufficient one.

Thus, what permits comparing the frequency (number) of anatomically identical responses is no different from what permits comparison in the case of anatomically disparate responses: average duration of the response events is the same. This point appears to have escaped the attention of several writers (e.g., Hodos & Valenstein, 1962) who have implicitly accepted either same-effect-upon-environment or anatomical likeness as a sufficient basis for comparing frequencies. The conclusion that in brain stimulation rate will not predict choice would appear to be the result of this malpractice, and probably can be rectified simply by restricting rate comparisons to events of comparable duration.

Moreover, response duration will provide a strict estimate of probability insofar as the distribution of responding is random throughout the considered interval of time (e.g., Parzen, 1960). Here we may recall the previously noted operating characteristics of the rat, the remarkable extent to which burst duration approximates a random process, which was so surprising. It is always gratifying to put surprises to work.

The above analysis would appear to justify the comparison

of anatomically unlike response events and thus would appear to establish the feasibility of testing the generalization. Tests were directed at four implications: relativity, nonuniqueness of consummatory events, reversibility, and indifference.

RELATIVITY

An original test was made with manipulation responses of Cebus monkeys (Premack, 1963a). By using manipulation—lever pressing of one kind or another—as both reinforcing and reinforced responses, the intent was to show that reinforcement could be produced without the use of goal responses. Tradition had divided the animal's repertoire into goal responses and non-goal responses or "instrumental activities" (the counterpart of neutral stimuli). Accordingly, the standard reinforcement paradigm was: make a goal response contingent upon an instrumental activity. Our gambit in this first test was thus to draw only from the nongoal category, and to demonstrate reinforcement notwithstanding, provided only that one nongoal response was more probable than the other.

Four manipulanda were used—a lever, plunger, hinged flap or door, and a horizontally operated lever—each of which could be used potentially both as the reinforcing or the reinforced response. A solenoid-operated lock made it possible to render any item inoperable; thus, the animal could be given any pair of items, with one of them locked and its operation made contingent upon the prior operation of the other member of the pair.

A two-stage test procedure was used. First, estimates were obtained of each animal's probability of operating each item. Second, the items were presented in pairs—one free, one locked—with operation of the locked item contingent upon operation of the free item. For example, the door and plunger might be given together, with the plunger free and the door locked. Release of the lock, and thus the opportunity to operate the door, was contingent upon the prior operation of the plunger.

Repeated measures on the items, presented one at a time with no response restrictions, showed the animals to differ substantially. Of four monkeys tested, one showed reliable differ-

ences between three of four items, another showed no reliable differences whatever, while two others were intermediate.

In a second procedure, still aimed at estimating response probabilities rather than testing the generalization, the monkeys were presented with *pairs* of items, again with no response restriction. The main purpose was to determine whether simply pairing the items would disturb the ordinal relations among the individual response probabilities. Inevitably, contingency tests will require pairing items; thus, if merely pairing the items should disturb the rank order of the individual response probabilities, tests of the individual items could not reasonably serve as a basis for predicting reinforcement values. Comparison of the outcome here with that of the previous procedure showed that (1) frequency of responding to any item is apparently always less when the item is paired with another (nonzero) item than when it is presented alone, but (2) the ordinal relations among items are not disturbed simply by the pairing procedure.

Contingencies were next arranged between the pairs of items, using a continuous reinforcement schedule and a rule that required at least one free response between any two contingent ones, thus precluding the hoarding of response possibilities. The six pairs of items (produced by the four manipulanda) generated twelve possible contingency pairs, since in this procedure each item can serve both as a free item on one occasion and as a contingent item on the next occasion. For each pair, a block of four daily contingency sessions were given, followed by a block of four daily extinction sessions, after which the animal was given a new pair of items and the procedure repeated.

The outcome for the four animals differed in a manner generally predictable from their original protocols. Results for Willy, the no-difference animal, were clear cut: rather than showing an increment on any of the five pairs on which he was tested, he declined on all pairs.⁵ Willy's overall decrement would

⁵ At least for the present set of items, Willy should be unreinforcible; it may be possible, by habituating him for some items and not for others, to establish the difference needed for reinforcement, but then the ceiling would be low indeed. Would a different set make Willy look like Chicko, or are the differences more a function of the organism than the items? An interesting

be expected on the following grounds: With no countering incremental effect of reinforcement, the 24-hr. interitem interval that was used in contingency training should produce a lower rate than the 96-hr. interitem interval that was used to obtain the original estimates of response probability (Premack & Bahwell, 1959). The contrast in the rest intervals used for contingency training and for estimation of response probabilities is pertinent to all the results. All increments will be conservative estimates: any increment produced by the contingencies will occur despite an interval that by itself would be predicted to produce a decrement.

The outcome for the two intermediate monkeys, Bimbo and Gimbel, showed one strong and two weak incremental effects; further, in neither *S* did any less probable response increase the frequency of any more probable response. But the clearest predictions possible were those for Chicko, who in the first procedure showed three reliably different response probabilities. Indeed, Chicko's protocol made possible three kinds of contingencies: contingent response higher than, less than, and, in one case, about equal to the free response. In brief, the outcome for the three types of contingencies were as follows: (1) contingent response higher than free response produced in all five cases an increment in the free response; (2) contingent less probable than free response produced in all three such cases a decrement in the free response, which is not predicted by the generalization—it says simply that such a case will not produce an increment—but which may be merely the result of the unfavorable intersession

individual difference would be the tendency to visit a fixed amount of relative preference among the items of any set. Some organisms may display preference among the items of all sets, while others may show no preferences among the items of any set. Thus Willy's flat profile might obtain generally, for all sets, in which case he should be insusceptible to reinforcement. This would be operationally equivalent to a defective reinforcement mechanism, in another idiom to "the salt has lost its savour," or in still a third idiom to the flat-affect model of Meehl's (1962) several models for schizophrenia. We might thus distinguish between two behavioral views of mental disease. In the traditional one, the reinforcement mechanism is intact but the history of reinforcement is untoward, whereas in another, the mechanism itself is defective, either genetically or as a result of a previously untoward reinforcement history.

interval; and (3) the one case in which the two responses were about equal produced little or no change, increment or decrement.

A last major point of interest in Chicko's data was the evidence concerning the relativity of reinforcement. H was the most probable member of Chicko's response set, P the least, and D intermediate. H reinforced all members of the set; P reinforced none (and instead was itself reinforced by all other members). Considered jointly, the results for H and P appear to substantiate the traditional absolute view: H, which is a reinforcer, reinforces all responses, while P, which is not a reinforcer, reinforces none. This appears to exemplify the "some are, some are not" view, as well as to support the standard trans-situationality assumption—H reinforced not one member of the set but all members. However, the results for D controvert the traditional view; D, which was of intermediate probability, reinforced P but failed to reinforce H. Thus, depending upon the relative probability of the base response, D both was and was not a reinforcer.

Since the present case appears to represent the first direct test of the trans-situationality assumption, the previous success of the assumption may rest upon a failure to have tested it. Indeed, we are now in a better position to contemplate the success of this assumption. The evidence for it is based entirely upon the demonstration that food (in the starved rat) reinforced all challengers—bar press, string pull, roll a marble. But I think we may see that, rather than proving trans-situationality, this more likely demonstrated that for the given parameters eating was simply the most probable member of the set; certainly all the events reinforced are low operant-level events, while eating in the starved rat can attain substantial probabilities. One response in any set would be predicted to be trans-situational: the most probable member of the set. In using eating and a starved rat, tradition worked exclusively with a most probable member. But a general theory of reinforcement cannot be based upon the special properties of the most probable member of the set. Reinforcement is not confined to this member, and the

properties of the other members differ from those of the most probable member. For example, they are not trans-situational.

REINFORCEMENT OF EATING AND DRINKING

Two further basic assumptions that go hand in hand are that consummatory responses are reinforcing and that reinforcement is a reversible relation. Food or water is customarily used to reinforce the bar press or running, but it is not asked: Can this relation be reversed? Will the bar press or running reinforce eating or drinking?

Are there intervals of time in which eating or drinking are less probable than certain other responses? If there are, then eating and drinking should be reinforceable. The present assumptions do not, of course, make any predictions of the above kind; they deal only with the effect of one response upon another, and thus can make no predictions until after the individual response probabilities are given.⁶

As to whether or not organisms could be found for which nonconsummatory responses take priority over consummatory, we first considered the need for an advanced organism—something pretty ethereal. The rat did not seem apt to produce so sublime an interval of time. Accordingly, the first test was made with children. Only later did we lift the ban on rats, having found appropriate intervals for them too.

In a first test using children, a pinball machine rewired for continuous operation and a candy dispenser, the two placed side by side, comprised the experimental arrangement (Premack, 1959). Candy consisted of constant-size chocolate bits, delivered one at a time by a conveyor belt into a dish each time the child ate the piece in the dish. Thirty-three children, the entire first-grade class of a public school, served as subjects. Their average age was 6.7 years.

On the first test, both the candy and the pinball machine were

⁶ We have not attempted to provide a general characterization, let alone theory, of events that for a given species lead to a high probability of responding, though such a theory could be of substantial value; e.g., it might be used to deduce the present reinforcement generalization. Cf. Pribram (1963) for an intriguing set of suggestions toward such a theory.

available without restriction. Sixty-one per cent of the children made more pinball-machine responses than they ate pieces of candy. (The average duration of the pinball-machine responses was just slightly greater than the visually approximated duration of the ingestion of a piece of candy. Thus the advantage in frequency of pinball responding would convert into an even greater advantage in terms of response duration; since the difference was not great, however, the scores were left in frequency as in the manipulation case.) We thenceforth referred to those who preferred the machine as "manipulators," the others as "eaters."⁷

On the second test, the availability of the candy and the pinball machine were made subject to either of two contingency relations, E-M or M-E. For E-M, each operation of the machine was contingent upon the prior ingestion of a piece of candy, whereas for M-E, each piece of candy was contingent upon the prior operation of the machine. Both the "manipulators" and the "eaters" were randomly divided, and half of each main group tested under E-M and half under M-E. For a manipulator, E-M was the experimental condition, M-E the control; the reverse held for eaters. What characterizes both experimental conditions is that the more probable response is contingent upon the less probable response (manipulators must eat to manipulate; eaters must manipulate to eat). Control conditions make the more probable response freely available, as in the first test, and thus provide a measure of changes in the less probable response that may occur independent of the contingency.

In both cases, making a more probable response contingent upon a less probable one increased the frequency of the latter. Thus, while candy reinforced manipulation in the eaters, operation of the pinball machine reinforced eating for the manipulators. Increments in the control condition were small compared

⁷ "Manipulator" is somewhat of a misnomer. Only some of those who preferred the machine emphasized the manner in which they shot the ball, shooting first from the hip, then with their back turned, all the while doing a kind of soft-shoe routine and practically never glancing at the ball. Others shot the ball with a minimal gesture, then feasted visually upon each new bank of lights the ball lighted up.

to those for the experimental condition, and the mean differences within both main groups were significant at less than the 1 per cent level.

The results indicate that the consummatory response is not unique in being a reinforcer of other responses but not reinforcing itself. Also, there would seem to be little doubt that with sufficient food deprivation, "manipulators" would become "eaters" and thus the response which had been instrumental could be made reinforcing and vice versa.

REVERSIBILITY OF THE REINFORCEMENT RELATION

Reversibility of the reinforcement relation was shown ultimately with the rat (Premack, 1962). With free access to both food and an activity wheel but access to water for only 1 hr./day, mean total drinking time for a group of six female rats was about 240 sec., while mean total running time in the same period was only about 1 min. On the other hand, with free access to both food and water but access to the wheel for only 1 hr./day, mean total drinking time per hour was now only about 28 sec., while mean total running time in the same period was 329 sec. These are by no means the only parameters that can be used to produce a reversal in the probabilities of drink and run. Combining sucrose concentrations appropriately with force requirements on the wheel will produce the same outcome. So deprivation is hardly the only effective parameter. But whatever parameters are used to reverse the response probabilities, the same prediction follows. It should be possible not only to reinforce drinking with running but also to reverse the reinforcement relation in the same subject simply by changing from one parameter value to another.

The apparatus used to test the predictions consisted of a modified Wahmann activity wheel equipped with a brake and a retractable drinkometer. The rat is placed in the wheel, which itself serves as the test cage. With drinking contingent upon running, the wheel was free but the tube was retracted; a predetermined number of runs brought the tube into the wheel for a predetermined time. Conversely, with running contingent upon drinking, the tube was present at all times but the wheel was locked;

a predetermined number of licks released the brake on the wheel for a predetermined time.

The more interesting case, run contingent upon drink, was tested first. Four female albino rats were given daily 1-hr. conditioning sessions, followed by daily 1-hr. extinction and reconditioning sessions. To augment the output, an FR was used in which each five licks freed the wheel for 10 sec. Throughout this training, food and water were continuously available in the home cage; after the last reconditioning session, water was removed from the home cage, and on the following day, training was begun with the reverse contingency—drinking contingent upon running.

With run contingent upon drink, total drink time was increased in all Ss by a factor of from three to five. For operant-level drinking, with only the tube present, mean total drink time was about 28 sec./hr., with both tube and wheel present about 23 sec./hr., and with run contingent upon drink about 98 sec./hr. The first extinction session further increased mean total drink time to about 175 sec./hr., following which drink returned essentially to its base. The picture was readily completed by changing the parameters so as to make drink more probable than run; then, with drink contingent upon run, the conventional increase in run was easily obtained. In summary, parameters were shown that made run more probable than drink and vice versa, and subsequently, it was possible not only to reinforce drink with run but also to reverse the reinforcement relation in the same subjects merely by changing from one set of parameters to the other.

INDIFFERENCE PRINCIPLE

Each time a response such as run is successfully predicted to reinforce drink or vice versa, a contribution is made to the indifference principle. This principle holds that prediction of reinforcement can be made without regard to the history of the response probability, for the outcome is indifferent to either the parameters used to produce the probability or the responses that manifest it. Successful prediction for such pairs as run and drink

would appear to bear this out, for there is hardly any doubt as to their being different neurophysiological systems. Indeed, the interest of the indifference principle accrues from this difference; the principle does not deny the neurophysiological difference, which indeed could not be denied, but asserts that despite the difference, the reinforcement value is predictable from a single scale of response probability. The principle says that the neurophysiological difference does not participate in or mediate the reinforcement value, which is determined solely by the duration for which the animal engages in the behavior.

Sharper tests of the point can be made by a somewhat different procedure. With a third response such as the bar press, which is less probable than both drink or run, the increment produced in the bar press by drink and run can be compared. If identical probabilities of run and drink are made contingent upon the bar press, the indifference principle requires that the increment to the bar press be identical. It further requires that if identical probabilities of drink—produced in one case by deprivation, in another by sucrose concentration, in still another by temperature or any other parameter—are made contingent upon the bar press, then, again, the increment to the bar press must be identical. Likewise, it requires that the same experiment in the case of running—where, in principle, identical probabilities of running can be produced by procedures ranging from oestrus to starvation—must also produce the same outcome, identical increments in the bar press.

Since identical response probabilities are easier described than produced, an approximation to this experiment was performed (Premack, 1963b). A set of response probabilities was produced that included at least one probability of running bounded by probabilities of drinking and at least one probability of drinking bounded by probabilities of running. When subsequently made contingent upon the bar press, the ordered set of run and drink probabilities should produce an identically ordered set of bar-press probabilities.

In Figure 1, bar presses per session are plotted as a function of the associated response probability. The points in the figure

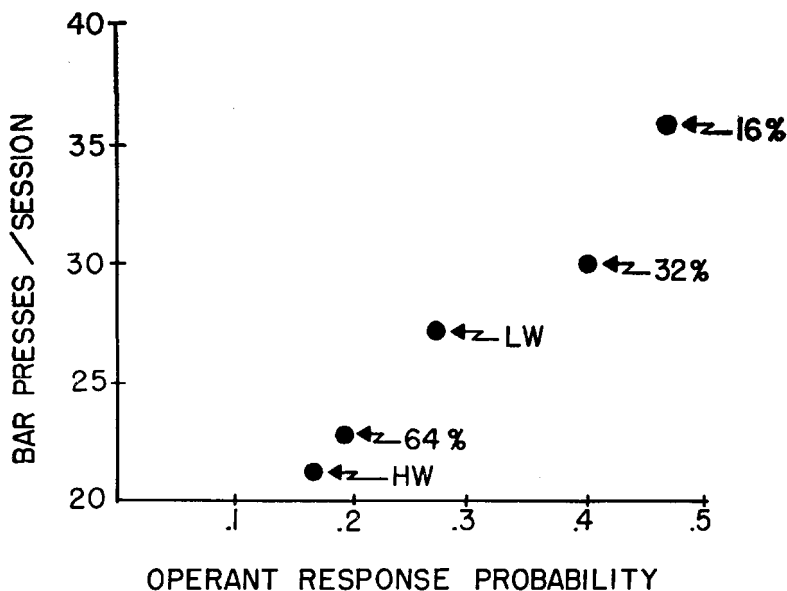


FIG. 1. Mean bar presses per session as a function of the associated operant-level response probability. The abscissa shows proportion of the operant-level session for which the animal responded—duration in seconds for which it ran or drank divided by duration of the session. Points are labeled according to the sucrose concentration (16, 32, and 64 per cent) or force requirement (light, LW; heavy, HW) that was used to control the operant-level response probability (after Premack, 1963b).

are labeled according to the sucrose concentration or force requirement that was used to produce the given drink probability in one case and run probability in the other. The functional relation shown there clearly supports the indifference principle: bar pressing increased monotonically with the associated response probability, and did so whether this was a probability of drink or run. These results were more encouraging than others that might be regarded as more dramatic. The universe hardly guarantees that comparable durations of run and drink will have comparable reinforcement effects; yet they clearly appear to.

The indifference principle is an exceptionally strong assumption, not to be established by any one test. If it should prove

to be valid for any large number of parameters, it would perhaps contribute more determinately to a unitary-state concept than any traditional drive notion. Are equal probabilities of eating, produced in one case by food deprivation, in another by sucrose concentration, and in still another by lowered temperature, nonetheless the same in their reinforcement value? An animal may be capable of differentiating states of hunger according to the specific antecedents of the hunger; yet it may not differentiate the reinforcement value of the states on any basis other than the associated probabilities of eating.

AN INCREMENT POSTULATE: THE EFFECT OF THE BASE RESPONSE

If we apply different more probable responses to a common base response, we may expect the increment to be proportional to the probability of the contingent response. A version of this experiment was shown in the application of varying probabilities of run and drink to the bar press; the greater these probabilities, the greater the increment to the bar press.

But consider the converse experiment. Suppose the same reinforcer is applied to base responses that differ in their probability of occurrence. Will the reinforcer bring all the base responses to the same level despite their pre-reinforcement differences? Or will the latter be reflected in asymptotic reinforced responding?

A first version of this experiment was done with Chicko via manipulation responding. H, the most probable member of Chicko's set, was applied successively to P, D, and L, responses less probable than H but quite different among themselves. H did not overcome the different operant levels of P, D, and L, but, rather, produced asymptotic rates that were proportional to the original operant levels. Thus, the operant levels showed through the reinforcement.

R. W. Schaeffer (in press) performed a more determinate version of this same experiment using groups of rats with drink and run as response materials. The experiment in the rat had been held back by the limited variation in probability of base response that is afforded by the traditional conditioning materials, viz., bar press and eating. But with the reinforcement of drink by

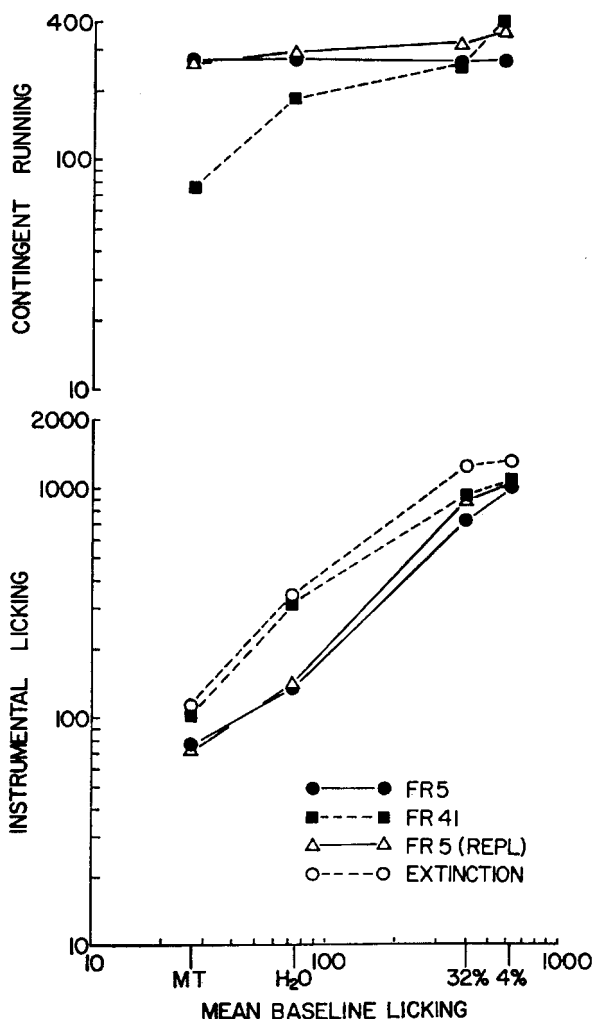


FIG. 2. Mean instrumental licks per session as a function of the mean operant-level licks per session in the operant-level session immediately preceding the contingency (after Schaeffer, in press).

run, Schaeffer was given nearly ideal materials. Run was made more probable than drink (mainly by giving only limited access to the wheel), and the probability of drinking was varied widely

over the four groups by tube content ranging from nothing to 32 per cent sucrose. After first establishing clear differences in the operant level of drinking, running was made contingent upon drinking in all four groups. Figure 2, which plots asymptotic licking as a function of the pre-reinforcement level of licking, shows that, as in the case of monkey manipulation, the same reinforcer does not bring base responses of different probabilities to the same level. Rather, the asymptotic response level was proportional to the pre-reinforcement level; once again, the operant level showed through the reinforcement.

A third version of this experiment has recently been performed by Holstein and Hundt (in press), again with rats, but now using brain-stimulation bar pressing and licking. Drinking a 25 per cent sucrose solution reinforced both brain-stimulation bar pressing and plain bar pressing, but for all rats the duration of bar pressing was greater the greater the pre-reinforcement duration.

REINFORCEMENT OF INTRACRANIAL SELF-STIMULATION

The position which eating or drinking once held has now been pretty much usurped by intracranial self-stimulation (ICS); it is the new king of reinforcers. The evidence leaves little doubt that ICS can be made extremely potent—the right combination of brain site and intensity of stimulation—but from the present point of view, intermediate strengths of ICS are of more interest than the maximal strengths. The use of only maximal strengths may lead to the same errors as did the use of only maximal or high strengths of eating. Indeed, this already appears to be the case; ICS-effective sites are described as the neurological substrate of reward, as nice an absolute theory of reinforcement, now in neurological garb, as could be imagined.

Electrical stimulation of the brain, like other kinds of stimulation, whether of taste or kinesthesia, for example, represents yet another means of producing response probabilities. Do response probabilities produced by ICS act like other response probabilities as regards reinforcement? If so, and a response can

be found more probable than the ICS-contingent response, then ICS should be subject to reinforcement.

Holstein and Hundt (in press) tested this matter by using drinking as the more probable response. They implanted three male Long Evans rats with bipolar electrodes, each at a different site to determine whether the obtained effect was circumscribed by site. Post-experimental histologies revealed the tip of the electrodes to be in the midbrain tegmentum, preoptic region, and median forebrain bundle, respectively. A constant-current 100-cps sine-wave generator was the stimulus source. Stimulation remained on for the duration of the bar press. Current was manipulated individually for each rat until each produced a moderate, stable response rate during daily session of 15 min. length. These current values were then maintained throughout the remainder of the experiment.

A retractable drinkometer, containing 25 per cent sucrose by weight, was introduced into a standard test cage, three inches to the side of the bar. Animals were given daily 15-min. sessions with the tube present and the ICS bar removed, followed immediately by another 15-min. session with the ICS bar present and the tube removed. The order of the two kinds of sessions was alternated from day to day. For all animals, the asymptotic duration of drinking proved to be greater than that of ICS bar pressing independent of the order of presentation.

A second pre-reinforcement measure was taken, as in the monkey manipulation case, by making the tube and bar concurrently available. Especially recommendatory of this step in the present case is the possibility of a recoil response to the ICS, which would not be recorded simply by a clock registering duration of the bar press. One approach, and nearly the ideal one, would be to arrange to include the recoil in the duration of the response. This did not prove to be necessary here. If recoil were present, it would compete with licking in the concurrent situation where both brain stimulation and tube were available and would reveal itself by an appreciable reduction in the duration of drinking; no such reduction was observed. Rather, the

ordinal relationship between the durations of drink and bar press were maintained relative to the individual presentation of the items, and in all cases the duration of drinking exceeded that of ICS bar press by a considerable margin.

A test was made by making access to the tube contingent upon ICS bar pressing. The bar was present at all times, but the tube was retracted and made available only after 10 bar presses. A clear-cut reinforcement effect was shown, the average duration of ICS bar pressing being increased by a factor of at least 2 by the contingency. Further, this increase was extinguished, that is, lost upon removal of the contingency.

A control was run on the unlikely possibility that brain stimulation could, without any change in parameters, reinforce drinking. Since we have attributed the reinforcement of ICS by licking to the relatively greater probability of licking, it would be highly embarrassing if, upon reversing the contingency, ICS should now reinforce licking. The contingency was reversed, and only one lick was required per ICS-bar press opportunity in light of the reported difficulty of generating high FR responding with ICS. No increase in licking was produced (by ICS bar pressing less probable than licking). An experiment in progress examines the effect upon licking of an ICS bar press more probable than licking, but the outcome does not seem especially uncertain.⁸

These first results suggest that response probabilities produced by brain stimulation do not differ from those produced otherwise. When the probability of ICS was greater than that of licking, ICS reinforced licking, but when the probabilities were reversed, licking reinforced ICS. If ICS is ascribed to a reinforcement center, we must now account for the fact that we can reinforce the events of that would-be center, and to do so we must find another center, and so on. Or if we place both events in the same center, we must then admit that events of the reinforcement center can themselves be reinforced as well as rein-

⁸ This prediction has since been confirmed by S. Holstein. In rats for which ICS was more probable than licking, making ICS contingent upon licking produced an increase in both the frequency and duration of licking. Personal communication, 1965.

force, and then it is doubtful whether we have any longer what tradition contemplates as a reinforcement center. Additionally, we must account for the fact that event A, which reinforced event B on one occasion, can be reinforced by B on another occasion. All of these difficulties can be resolved by ceasing to regard reinforcement as an absolute property, and thus ceasing to suppose that a center exists which supports all reinforcement. A more reasonable anticipation may be a mechanism for sensing the dominance of one event relative to another, along with the predilection of the lesser event to adopt or conform to the properties of the greater event. What have so far been called "reinforcement centers" would more nearly appear to be locations that support dominant, i.e., high-probability, events.

CHOICE

The ordering a subject gives to a set of stimuli can operationally define the basic motivational fact of *preference*. We have already looked at three procedures for measuring preference, and in turning now to choice we consider still another. In the monkey manipulation study, you will recall, a set of items were ordered in three ways: first, by the total duration for which a subject responded to the individual items; second, by which item would reinforce (responding) to which other items; and third, by the increment an item produced in a common base response when made contingent upon that response. Although the monkey data were only partially confirming, we nonetheless assume that the three procedures produce equivalent orderings: if S responds longer to A than to B, then A contingent upon B will reinforce B, but not vice versa; and A contingent upon X will produce a greater increment in X than B contingent upon X. Can choice now be added to this list of assumedly conceptually equivalent procedures? It would seem reasonable to suppose that it could. But this is rather a weak assumption, less interesting than others that could be made. A more important question is whether we can find any formal overlap or commonality between choice and any of these other ordering procedures. If we can, then quantitative relations that have already been established for the case

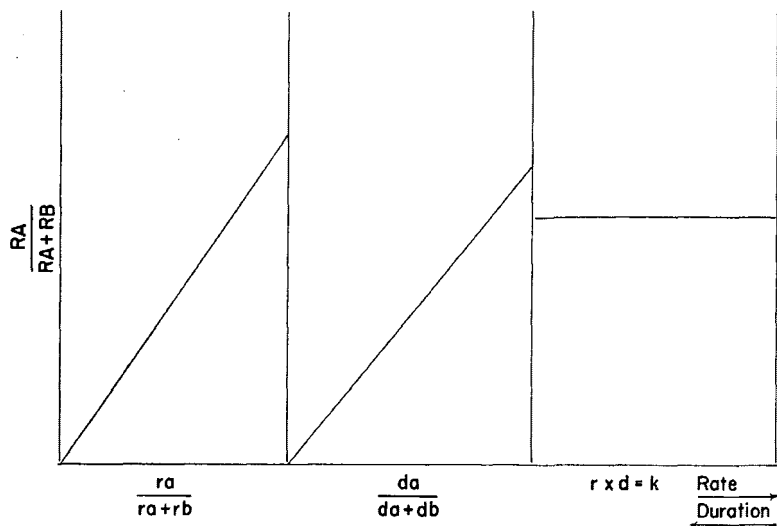


FIG. 3. Relative rate of responding on bar A, $\frac{RA}{RA + RB}$, as a function of (1) relative rate of reinforcement, $\frac{ra}{ra + rb}$, (2) relative duration of reinforcement, $\frac{da}{da + db}$, and (3) relative total duration of reinforcement, $r \times d$. See text for further explanation.

of choice can be extended to the other preference-measuring procedures.

Important data for the free-responding choice situation have been provided by Herrnstein (1961), Reynolds (1963), and Catania (1963a). Using hungry pigeons in a two-key situation with grain available on both keys according to VI schedules, these workers have found the birds to distribute their responding to the two keys in an impressively simple manner. The proportion of the total responses the bird makes on a given key is equal to the proportion of the total reinforcements it receives on that key. This relation is shown in Figure 3, left panel, where relative rate of responding on A, $\frac{RA}{RA + RB}$, is plotted as a function of the relative rate of reinforcement on A, $\frac{ra}{ra + rb}$.

In Figure 3, middle panel, we have extended this same kind of relationship to another parameter, viz., reinforcement duration, although the actual evidence for this second case is slight (Catania, 1963b). Nevertheless, it seems only moderately risky to consider that if, rather than varying the schedule on the two keys (while equating the magazine duration), the magazine durations were varied (while equating the schedules), one would obtain the outcome shown in Figure 3. Let us make the two cases more explicit. In the first case, duration for which the grain hopper is available following each completion of the instrumental requirement is the same for both keys, while rate of reinforcement is varied; whereas in the second, rate of reinforcement is the same for both keys, while duration of reinforcement is varied. We assume that the two parameters yield identical relations.

If, in fact, rate and duration prove to yield comparable functions, it would be additionally desirable to reduce both cases to the same variable. A variable that would have this effect is simply *total reinforcement time*, i.e., the rate by duration product. This product is necessarily varied by experiments which hold reinforcement duration on the two keys constant while varying rate, no less than by those that hold rate constant while varying duration of the individual reinforcement on the two keys. That is, these two kinds of experiments simply represent two different ways of randomly distributing a block of reinforcement time (more probable responding) over some larger interval of time—in one case, by number of instances, and in the other, by duration of the instances. Now the rat has only to respond indifferently to variation in number and duration in order to make a neat, if small, contribution to parsimony; we are testing the matter now.

In progress for some time have been experiments in which equal total duration of reinforcement, produced by different combinations of rate and duration, are associated with the two bars in a two-bar choice situation. For example, associated with one bar are 30 4-sec. opportunities to drink water, and with the other bar, 60 2-sec. opportunities of the same kind. The pre-

dicted results are shown in Figure 3, right panel; the rate and duration scales on the abscissa have been pitted against each other in such a manner as to yield equal $r \times d$ products at all points; accordingly, responding is predicted to be equally divided between the two bars at all points. It seems little to ask, so little that after a year of technical difficulties our position on the matter has become: if the rat declines this invitation to parsimony, it is an ungrateful creature and not to be trusted further.

What relation, if any, can we now find between this proposed extension of the work on schedules and the other preference-measuring procedures that we have already considered? Since we have already indicated how it may be possible to change the rate measure, in which schedule problems are typically formulated, into a duration measure, the reader may wish to complete the parallel for himself. Perhaps, however, in order to assure that all parties complete the parallel in approximately the same way, it may be advisable to make our version explicit.

The transition from choice to the other procedures can be made via two further points. The first concerns the proportion of magazine time for which the birds actually eat in the Herrnstein, Reynolds, and Catania work. These durations have not been reported, but because of the parameters that have been used, it is reasonable to assume that the durations are essentially constant. That is, it is reasonable to assume that, first, the proportion of time the birds actually eat from one magazine presentation to the next is stable over the session, and second, this proportion is constant over the range of variables that have been tested. For example, these workers have not combined weak hunger with "long" magazine duration, which would be expected to lead to a within-session decrement (see Collier & Siskel, 1959; Collier, 1962), nor have they combined "short" magazine durations with "long" intermagazine intervals, which we have reason to suppose would substantially reduce the proportion of the possible magazine time that was utilized. In brief, though not commenting on the fact, these workers have restricted themselves to a narrow range of parameters—a range which makes the

assumptions concerning constancy of the utilized proportion of magazine time quite reasonable. These expectations are supported to some extent by the report that the birds take all of the reinforcements that are offered by the schedules.

Now our concern with *actual* eating time, as opposed to the experimenter's parameters of rate and duration, which merely define a *possible* eating time, is to complete the transition from the variables in which the problem has been stated to the variable we consider to be critical (here no less than in the cases we have already examined). For this case, as for the others, we would make the same assumptions: (1) eating reinforces key pecking in the starved pigeon because eating is more probable than key pecking (this point is worth making because the tendency to speak of reinforcers as stimuli drawn from a special drawer marked "reinforcer" continues unabated), and (2) the relative reinforcement value of eating is equal to the relative total duration of eating.

That is, we expect possible reinforcement time to predict instrumental rate only when possible proves to be an unbiased estimate of actual. If, for example, the work on choice were extended to parameters beyond the narrow range used so far, we would *not* expect relative rate of responding to be a simple function of either relative rate or relative duration of reinforcement. Thus, in our present experiments, where the opportunity to drink water is used as the reinforcer, we do not expect 80 1-sec. opportunities to have the same effect as 20 4-sec. opportunities. Measures of drinking taken both within and outside the reinforcement situation show the proportion of time the rat drinks to be less with a 1-sec. magazine time than with a 4-sec. time. When actual magazine response time is *not* a fixed proportion of possible magazine time, we would predict the simple relations shown thus far to vanish. But we would expect these simple relations to be promptly restored, for all parameters, simply by plotting instrumental responding as a function of actual rather than possible magazine response time.

Turn now to the last point upon which completion of the

parallel depends. This concerns the manner in which the experimenter determines the proportion of time an organism spends responding to one of several alternatives. The reader may have observed that this proportion is determined more directly in the pigeon experiments than in the experiments considered earlier. In the former, the bird cannot eat more than, say, 20 per cent of the total eating time in association with one key for the simple reason that 20 per cent is all the experimenter provided on that side. In the experiments considered earlier, however, the animal could respond in all possible proportions to either alternative, for this was not limited by the experiment. It was not limited directly, at least, as in the pigeon experiments, but only indirectly by stimulus and/or deprivation factors. Although the animal was operated upon indirectly by being given particular stimuli to choose between, or being allowed to choose after varying deprivation periods, it could nonetheless divide its responding in all possible proportions rather than according to limits set by the experimenter.

In the one case, the experimenter bakes a pie, smaller than the bird can eat, and divides it making, say, 20 per cent available on one side. By making the pie smaller than the bird can eat (and serving it in small pieces at suitable intervals), he assures that the momentary probability of eating remains constant, and thus assures that the bird makes good on the 20:80 split. This represents a direct way of predetermining the distribution of the more probable response event, in this case eating, over the several alternatives. It may be seen to contrast with the relative indirection of giving the subject two pies, both larger than it could eat, and measuring the duration for which it eats one or the other.

Nevertheless, both the direct and indirect procedures affect the same variable, viz., the distribution of the more probable response events over the alternatives. Consequently, we are obliged to predict that their effects will be the same. That is, unless we introduce special assumptions, we are obliged to predict that whether the distribution is forced upon the animal by

rate and/or duration parameters or generated by the animal's own dispositions, the effect upon instrumental responding will be the same. That is,

$$\frac{P(RA)}{P(RB)} = \frac{cFa}{cFb} = \frac{T_a}{T_b}, \quad \text{where} \quad \frac{T_a}{T_b} = \frac{\frac{t_a}{t_a + t_b + k}}{\frac{t_b}{t_a + t_b + k}}.$$

$P(RA)$ is probability of A instrumental responses, c is a constant determined by magazine duration, F is frequency of magazine presentation, t_a is duration of S 's response to item a , and k is a constant equal to session time minus t_a plus t_b . An experimental example may be clarifying.

Picture a Skinner box temporarily bereft of levers but containing both a right and left magazine (or one magazine sometimes accompanied by a red light, other times by a green light). The two magazines are operated successively, each according to some predetermined schedule. The organism is hungry beyond the possibility of satiating on what the magazine opportunities provide. Measurements are made of the duration to which S responds to each of the magazine opportunities; the Skinner box regains its levers, and the subject is advanced to the second step of the experiment, where it is obliged to respond instrumentally for the same magazine opportunities it experienced without an instrumental requirement at step 1. For some parameters, T_a/T_b will simply equal F_a/F_b , the ratio of the reinforcement schedules, but whether or not this special condition obtains, we predict

$$\frac{P(RA)}{P(RB)} = T_a/T_b.$$

In a second experiment, the Skinner box, again bereft of levers, offers the subject a choice between any pair of stimuli to which the species responds, e.g., run and drink, manipulate and eat, eat a on one side and eat b on the other. Here, too, the duration for which S responds to the alternatives is measured, and the experiment is then concluded as before. Although T_a/T_b is now determined by subject disposition rather than by experimenter

limit, we again predict $\frac{P(RA)}{P(RB)} = T_a/T_b$.⁹

It should not be thought that the two experiments differ importantly in that the stimuli are available discontinuously in one case and continuously in the other. This difference is trivial. The alternatives can be presented discontinuously no less in the one design than in the reinforcement-schedules case, though in the latter, discontinuous presentation is a necessity rather than an option. Custom presents stimuli continuously in the operant-level session. But stimuli need not be present throughout the session in order that their availability to the organism be noncontingent, i.e., independent of an instrumental requirement. The operant level for a discontinuously available stimulus can be used to predict the reinforcement value of the stimulus (relative to some other stimulus) no less well than can the operant level for the typical continuously present stimulus. Indeed, the former will afford a somewhat more accurate prediction than the latter. The very procedure of reinforcement imposes a discontinuity upon the availability of the contingent stimulus, a discontinuity which may affect the animal's probability of responding to the stimulus. The correction here is possibly small, a nicety at the present stage, though data can be shown in support of the possibility. Figure 4 shows amount drunk per session for three independent groups of thirsty rats that encountered a drinking tube which was (1) continuously available for 20 min., (2) in and out of the test cage alternately for 7 and 21 sec., respectively, and (3) in and out of the test cage alternately for 28 and 5 sec., respectively. The total period of availability (20 min.) was the same for all three groups; nonetheless, as Figure 4 shows, the groups drank different amounts ($P < .01$). Interestingly, discontinuous presentations were found that both increased and decreased total intake relative to that for the continuous stimulus.¹⁰

⁹ This prediction obviously cannot require that responding be constant over the considered interval, but it can and does require that the decrements in responding to the alternatives offered occur at approximately comparable rates. Ways of coping with responses that decrement or habituate at different rates have been suggested in Premack (1961) and, more generally, in Hundt (1964).

¹⁰ These data were collected by S. Manaster in 1961.

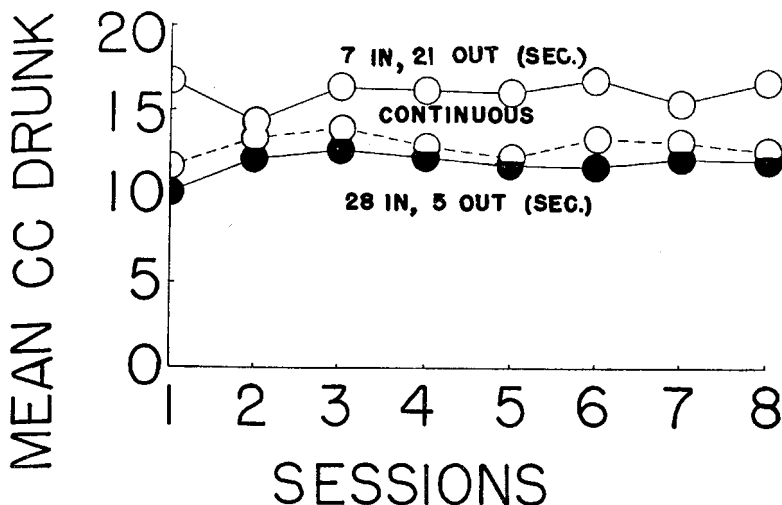


FIG. 4. Total amount drunk per session by each of three groups of rats that differed as to the manner in which the drinking tube was available to them, e.g., continuous *versus* discontinuous. See text for further explanation.

It is clear that the reinforcement schedule in the conventional problem will affect the distribution of the availability of the contingent stimulus. This effect will be direct in the interval schedule, more so than in the ratio schedule, though whether direct or not, all reinforcement schedules will have some effect upon how the availability of the contingent stimulus is distributed. What effect will the reinforcement schedule have upon the reinforcement value of the contingent stimulus? Perhaps no effect or perhaps a large one, in any case an amount that will depend, according to the present account, upon the degree to which the reinforcement schedule affects the animal's probability of responding to the stimulus. That is, if the schedule affects the animal's probability of responding to the stimulus (as in Figure 4), it will likewise affect the reinforcement value of the stimulus; otherwise it will have no effect. The time distribution of the stimulus thus joins other stimulus parameters, e.g., sucrose concentration, in being a parameter of reinforcement only insofar as the stimulus property is a parameter of

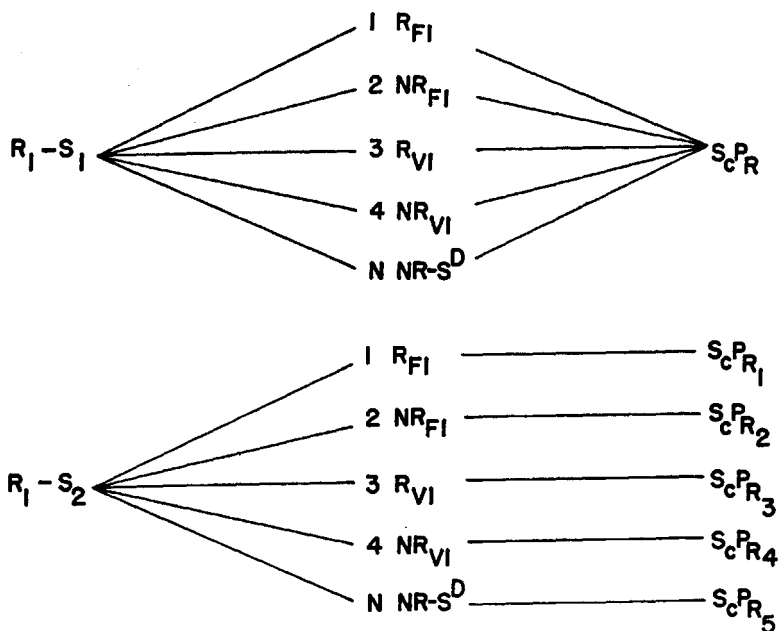


FIG. 5. Discriminative stimuli, S_1 (upper panel) and S_2 (lower panel), are associated with a common event, S_c , to which the species responds. The nature of the association varies, in one case affecting (lower panel) and in the other not affecting S 's total duration of responding to S_c . The prediction is made that the acquired reinforcement value of S_2 will vary over conditions, while that of S_1 will be constant over conditions.

response probability. This point is brought out graphically in Figure 5.

The upper portion of Figure 5 shows a discriminative stimulus, S_1 , to be associated with S_c , any event to which the species responds, e.g., mate, wheel, food, manipulandum. The nature of the subject's access to S_c varies markedly. In two cases, R_{FI} and R_{VI} , access to S_c depends upon instrumental responses that pay off according to fixed- and variable-interval schedules, respectively. In two other cases, NR_{FI} and NR_{VI} , S_c is independent of an instrumental requirement, being programmed, like the drinking tube in the experiment above, according to fixed or variable schedules. Finally, in $NR-S^D$ a signal is added indicating the

advent of S_c . The assumption depicted in the top half of Figure 5 is that none of these differences in manner of access yield differences in total duration of S 's response to S_c ; the bottom half of the figure depicts the opposite assumption. The prediction then made is that the reinforcement value of the discriminative stimulus will be constant over conditions in the first case, and variable in the second. Data bearing suggestively on this prediction have been reported (Catania, 1963a). In a choice situation, the rate at which pigeons pecked on key_A was shown to depend, not upon rate of responding to key_B, but solely upon rate of reinforcement received on key_B.

NEGATIVE REINFORCEMENT

Our basic measure for predicting positive reinforcement has been the duration for which a subject puts itself into a given state, relative to the duration for which it puts itself into some other state. It would seem reasonable to use the complement of this measure as a basis for predicting negative reinforcement, viz., duration for which a subject removes itself from a given state (into which the experimenter has forced it), relative to the duration for which it removes itself from some other state. Unfortunately, we have collected no relevant data, though a former student, A. G. Hundt, has formulated what would seem to be a basic question for this position. If the probability of S 's remaining in one state is numerically equal to the probability of S 's removing itself from another state, are the reinforcement values of the two responses equal? The present position requires an affirmative answer, and the question is now being tested with the use of food and electric shock, respectively.

Although, typically, negative reinforcement looks only at the offset of an event and positive reinforcement only at the onset, it should be observed that organisms both turn on and turn off virtually all stimuli to which they respond at all. Animals not only start eating but obviously also stop; as was clear in the *ad lib* case, eating is highly discontinuous, occurring in bursts and pauses. Although it apparently took intracranial self-stimulation (ICS) to call attention to the fact of on-off (Roberts, 1958;

Bower & Miller, 1958), it should now be evident that on-off is not a unique property of ICS, but rather is a general property of free responding. This leads to a further conclusion as regards the reinforcement possibilities for any stimulus to which the species responds; it should be possible to use the same stimulus both as a positive and as a negative reinforcer.

To demonstrate that the same event can be both positive and negative reinforcer, we require, first, an arrangement in which the onset of the event be more probable than the onset of some other event, and second, that the subsequent offset of the event be more probable than the onset of some other event. In the consummatory cases, this kind of arrangement could doubtless be made by the appropriate surgical preparation, for example, by tubing liquid directly into the animal. We have used instead a locomotor case, which retains the advantage of an intact animal. Two findings led to the arrangement: (1) for the rat, the opportunity to force itself to run is reinforcing, and (2) the rat is able to drink while running.

The rat is placed in a modified wheel containing a bar and a drinkometer (Hundt & Premack, 1963). The wheel is not free to move but is connected to a motor which the rat can turn on by pressing the bar. When the rat presses the bar, the motor is activated, the wheel rotates, and the rat is forced to run—has forced itself to run. It must continue running until it licks the drinkometer some predetermined number of times, which turns off the motor, stopping the wheel and allowing the rat to stop. The rat thus both forces itself to start and stop running by bar pressing and licking, respectively.

In one such experiment, an additional question was answered by using a fixed ratio of from 1 to 19 licks on the off response, viz., how does the "difficulty" of turning off a response affect the likelihood of its being turned on? Figure 6 shows the main results for a representative subject. Shown as a function of the fixed-ratio requirement on the off response are (1) frequency of the on response, (2) duration of the off response, and (3) duration of running. It may be seen that forced running increased bar pressing and that the frequency of the bar press (on response)

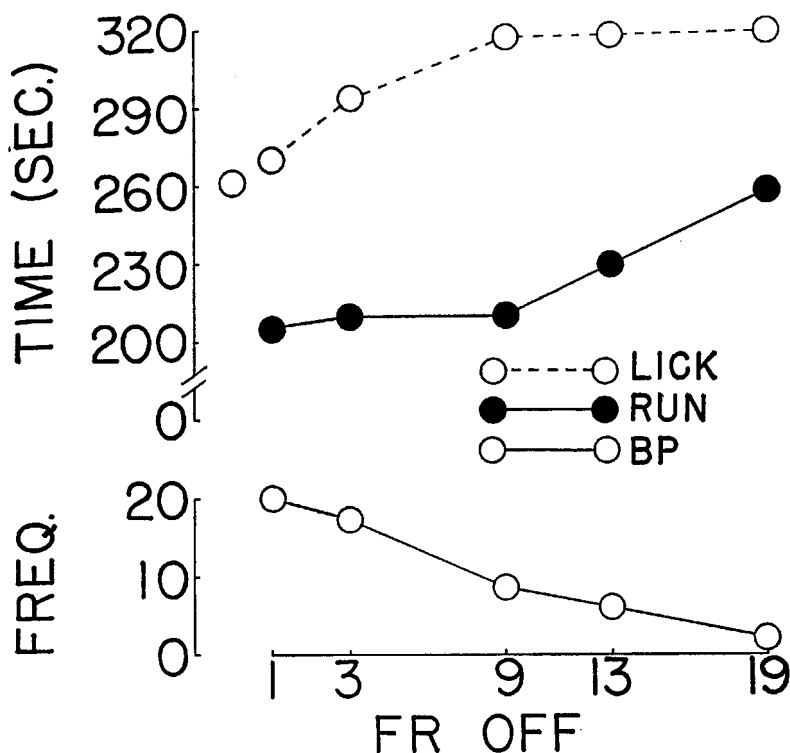


FIG. 6. Shown as a function of the number of licks required to turn off the wheel (FR OFF) are (1) average bar presses per session, (2) average duration of licking per session, and (3) average duration of running per session. The point to the left of the lick curve gives the base duration of licking (after Hundt & Premack, 1963).

was inversely proportional to the FR of the off response. The rat turned the wheel on only about twice per 20-min. session when 19 licks were required to turn it off, but turned it on upwards of 20 times per session when only one lick sufficed to turn it off. The average duration of licking and of running are both shown to increase over the FR variable, that of licking beyond the base duration (shown by the unconnected circle), so that we may conclude that the offset of running reinforced licking quite as the onset reinforced bar pressing. The total duration

of running per session increased over the FR variable despite the fact that the number of times the rat turned the wheel on decreased over the same variable. The increase resulted from a marked increase in the average duration of the burst; though the rat started running less often, once started, it continued for notably longer—e.g., average burst duration for FR 19 was 140 sec., compared to 10 sec. for FR 1. This did not result from the rat's trying to stop but being unable to do so because of the large lick requirement: the *first* lick did not occur for about 134 sec. after the animal started running in the FR 19 case, as compared to about 10 sec. for the FR 1 case.

These results indicate how a generalization that was stated originally for positive reinforcement may now be broadened to include negative reinforcement as well. Originally, the generalization read: for any pair of responses, the more probable one will reinforce the less probable one. But this fails to distinguish between the onset and offset of an event. The generalization should now read: if the onset or offset of one response is more probable than the onset or offset of another, the former will reinforce the latter—positively if the superiority is for "on" probability and negatively if for the "off" probability.

ACQUISITION

So far we have looked only at asymptotic responding, giving no attention to the manner in which reinforcement is acquired. Tolman¹¹ once remarked of the Skinner box that its data were of relatively little interest to him in that the learning that takes place there is over after the first or second bar press. The rest, he suggested, is performance and thus of less than vital interest to him. This may be seen as a comment on the efficiency of the apparatus. On the other hand, Tolman's comment may be viewed as a perceptive complaint—as noting that the steps taken to make the organism susceptible to reinforcement, to prep it, as it were, are buried in the artistry of the laboratory; these steps need instead to be made an explicit part of theory.

Consider that we have two events, one of them incontro-

¹¹ Personal communication, 1960.

vertibly reinforcing (H) and the other incontrovertibly reinforcing (L). What is the minimal relation that can obtain between H and L that will produce reinforcement? How slightly can they contact each another and the contact still be effective? This is the question we want to examine now, as opposed to the matter of what are reinforcing and reinforcing events. We have already urged that the latter consists of any two responses that differ sufficiently in their probabilities of occurrence. Let us suspend debate on that topic in order to ask: What is the weakest relation that more and less probable responses can have that will eventuate in reinforcement?

We all know one sure-fire procedure for producing reinforcement: the contingency, the "if L, then H" operation which is used to produce virtually all reinforcement. But is the contingency a minimal or even a weak operation? Does it provide only temporal contiguity between H and L—the incontrovertibly reinforcing and reinforcing events—and nothing else?

Consider an ordered set of operations, all of which may conceivably produce reinforcement. The weakest will be pure temporal contiguity between L and H, while the strongest we need consider here will be the contingency. There are more than a few ways to produce temporal contiguity between L and H, and these vary in strength according to the number and perhaps kind of additional conditions they superimpose upon temporal contiguity. Pure temporal contiguity, we will show in a moment, is rare or unknown.

In order to locate the weakest possible effective operation, we will start with the strongest operation in the set—which will give us the assurance that reinforcement is possible with the materials in question—and progressively denude the strongest operation of its surplus conditions until a point is reached where reinforcement is no longer produced. This point may be on either side of temporal contiguity.

What are the major surplus conditions of the contingency? Embodied in the contingency are at least three first-order conditions over and above simple temporal contiguity, not to mention some second-order conditions that are less evident though

possibly important. Consider the first-order properties more or less in the order of their visibility. *First*, the heart of a contingency, which is a response requirement: the reinforcing event can occur only if L occurs first; this is a special constraint, for if a contingency did not obtain, the animal might enter H from any of a number of antecedents. But a typical contingency prescribes that H can be reached only through a preselected L. *Second*, the distribution of the organism's responding to H is circumscribed. The animal has a characteristic way of responding to H, both in terms of burst durations and intervals between bursts, but a contingency limits both maximum burst duration and minimum interburst interval; it sets these limits dramatically when the schedule is intermittent, though even to some extent when the schedule is continuous. That is, by its very nature a contingency involves intruding an experimenter-selected event, L, between the organism's successive responses to H, thereby disturbing to one degree or another the organism's response distribution on H. *Third*, a typical contingency leads to a reduction in the total amount or duration of responding to H, relative to what would occur were H free. The amount of food the animal eats, for example, in a typical contingency is substantially less than the amount it would eat were food freely available for a period of time equal to that of the contingency session. This reduction is not a necessary consequence of a contingency; as we shall see, a contingency can be arranged that leads to no such reduction, but historically the reduction appears to be an invariant concomitant of the contingency procedure, and thus it remains to be seen whether the reduction plays a role in reinforcement. To be sure, any of the constraints described above may be as inconsequential as they appear to be accidental; the mere fact that we find them in the test tube does not prove that they are active ingredients; they may be inert accompaniments and temporal contiguity the only active factor.

The three above are simply the most visible surplus conditions of the contingency; we may mention at least one of the less evident possibilities. The response requirement, which allows H to be reached only through L, may be conducive to "errors"

or nonreinforcement, and these may be potentiating. That is, if L is not high in the animal's hierarchy, other responses may precede L and of course fail. These failures may call attention to the effectiveness of L. Speaking loosely, except as I have first failed on x, y, and z, I may not recognize the effectiveness of L when at last H is obtained through L. However unlikely a non-cognition theorist may consider such a mechanism, it should at least be recognized that its occurrence is more or less built into the contingency procedure.

This characterization of the contingency is intended to establish what is already suspected, viz., that the contingency is not a simple procedure, certainly not one whose effectiveness permits citing temporal contiguity as a sufficient condition. We know that the contingency works, but, in view of all that it contains, we hardly know why. We need to begin peeling off the surplus conditions.

A procedure which is weaker than the contingency and which is said to demonstrate that temporal contiguity is sufficient, is Skinner's (1948) noncontingency or superstition case. This procedure differs from a contingency principally in that it drops the response requirement. As you will likely recall, grain was available to pigeons in Skinner's classical demonstration according to a temporal schedule that was independent of the bird's responding. Since this procedure does not prescribe the path through which H is reached, the animal could enter H by any of a number of antecedents, different from bird to bird. Skinner reported pigeons to be reinforced by this procedure, to increase in the frequency of turning about or pecking the ceiling—any of a number of idiosyncratic responses which, though not required, occurred prior to H. Skinner concluded that temporal contiguity is a sufficient condition for reinforcement.

But we may note that although Skinner's procedure is weaker than the contingency, it is not yet a pure case of temporal contiguity. On the contrary, it retains from the surplus conditions of the contingency both a circumscription on the distribution of responding to H and a reduction in the total amount of responding to H. That is, the grain was not free but was available to the

birds according to some program, and the amount they ate was almost certainly markedly less than they would have eaten in an equal period of time had grain been freely available.

Can temporal contiguity be achieved without either a response requirement or any restriction on H? The free pairing of stimuli to which the subject has different probabilities of responding would appear to be one way. For example, suppose a rat is given unrestricted access to both an activity wheel and a source of water. Previous tests have established that with suitable water deprivation, the total duration for which the rat drinks when only water is available is greater than the total duration for which it runs when only the wheel is available. Accordingly, in this case, drinking will instance H, running will instance L.

Now when the rat is given both the tube and wheel at the same time, it will occasionally run and then drink—in the case of some of our rats, repeating this run-drink sequence perhaps 20 or more times in a 15-min. session. There will thus occur a temporal contiguity between incontrovertibly reinforcing and reinforcing events. That is, if, with identical parameters, drinking were made *contingent* upon running, running would increase without question. The temporal contiguity that interests us now, however, is one that occurs without the encumbrances of the contingency. If there is no response requirement, no circumscription upon the distribution of H responding, and no reduction in the amount of H responding, is temporal contiguity between H and L still a sufficient condition?

All the evidence we now have on this question is negative. Pure temporal contiguity has never proved to be a sufficient condition. By now a fair number of rats, upwards of 40, have been placed in a wheel containing a drinkometer and their running and drinking recorded on Esterline Angus tape. Despite the relatively large number of times these records show an instance of drink to follow (within 2 sec. or less) an instance of run, there are no cases in which the amount of running has increased over the base amount—the amount run when only the wheel is available; in fact, the rat typically runs more *before* the tube is inserted. Table 2 affords some idea of the number of times the

TABLE 2
NUMBER OF NONCONTINGENT RUN-DRINK CONTIGUITIES/15 MIN.

<i>Subject</i>	1.	25	7.	19
	2.	24	8.	17
	3.	22	9.	13
	4.	11	10.	23
	5.	13	11.	21
	6.	8	12.	23

would-be effective temporal contiguities occur in this situation. This number of contiguities in a contingency situation would be more than sufficient to produce an increase in the frequency of running. Further, the number shown is an average per session, and many sessions are run before a negative outcome is accepted.

It might be considered that when the rat can both drink and run and drinking is more probable, running fails to increase simply because drink suppresses run. That is, although run is reinforced by drink, the increased disposition to run is not manifested because of response competition. This would make more sense if the total duration of run plus drink approximated total session time, which it does not; even so, since a direct test can be made so simply, the point is worth considering. The possible suppressive effect of drinking can be removed altogether simply by removing the tube. When this is done, the animal being placed in the wheel with the drinkometer removed, the amount it runs is not different from the base amount. Thus, no increase in running is seen either during a potential learning situation or subsequently during a potential extinction situation.

Thus far we have described three temporal contiguity-producing operations, ranging from the strong contingency to the weak pairing of two stimuli (wheel and tube); intermediate is Skinner's superstition paradigm, which appears to differ from the contingency only in dropping the response requirement. Since the latter produces reinforcement and the free-pairing procedure does not, this leaves in doubt the effect of the other two principal encumbrances of the contingency—circumscription of the H distribution and reduction of H total. Both are retained in the superstition paradigm, which is effective, and both are

dropped in the free-pairing design, which is not effective. Accordingly, at least one and perhaps both of these conditions are necessary, i.e., must accompany temporal contiguity in order for reinforcement to occur.

One way to proceed now is to attempt to pervert the purpose of the contingency. Normally, a contingency is used to produce reinforcement, but is it possible to establish a contingency relation and yet not produce reinforcement? The instructive consequence of attempting to divert the contingency from its usual end is that it leads to the abandonment of one of the normal surplus conditions of the contingency—specifically, the reduction in H. Although a circumscription on the distribution of H is all but inevitable, a result of the instrumental requirement, contingencies can be formed that do *not* reduce the animal's normal amount of responding to H. Indeed, the characteristic reduction in H is a coincidental by-product of another accidental feature of the standard contingency. The typical instrumental response, e.g., bar press, has a low operant level, and when access to H is made contingent upon such an instrumental event, a reduction in H, particularly when the contingency is first instated, is nearly inevitable. For example, when the rat can obtain a pellet only following a bar press, the low operant level of the bar press will almost certainly reduce the number of pellets the rat eats in the beginning (whether the reduction obtains after bar pressing has increased in frequency is immaterial; at that point reinforcement has taken place, and we are concerned with whether the initial decrement in H played any role in the ultimate increment in L).

The reduction in H can be avoided from the outset, however, by using an atypical instrumental response, one that has a substantial operant level. For example, consider drinking to be made contingent upon running at a time when running has a high pre-reinforcement level, a thing which is easily established in the rat. By requiring only a "little" run for each relatively "large" drink, the arrangement can be made whereby the rat can provide itself the opportunity to drink the normal amount by running no more than the normal (base) amount. Now the

striking thing about this kind of contingency is that it produces no reinforcement, no increase in the frequency of running. But, indeed, why should it? Without running more than it would otherwise, the animal can (and does) drink essentially the normal amount. We are thus led to see that a contingency per se is not a sufficient condition for reinforcement. Moreover, this completes the devaluation of the response requirement: Skinner's superstition paradigm showed that it is not a necessary condition for reinforcement, and the present design shows that it is not a sufficient condition.

More instructive is the suggestion that, specifically, a reduction in H is needed in addition to temporal contiguity if reinforcement is to occur. Circumscribing the distribution of H is clearly not a sufficient condition (though it may be necessary), for the contingency will have this effect; yet if, despite the partitioning of H, H occurs in normal amount, there is no increase in instrumental responding.

But a main point may be considered to remain: if the rat drinks its normal amount, there is no reason for it to run more than its normal amount, but what if the tube is removed? Here, too, there is no increase. After numerous sessions in which instances of run led repeatedly to instances of drink, but with no decrease in drink and no increase in run, the drinkometer was removed, as in extinction, and the animal given only the opportunity to run. No increase in running was observed in three such studies involving 19 rats.

The failure of this kind of contingency to increase instrumental responding, during either a potential learning situation or, more strikingly, a would-be extinction situation, is illuminating on still another score. Data from these contingencies can be shown essentially to rule out the possibility that, during the contingency, runs occur which do not lead to drink and thus are extinguished. We have not mentioned this point earlier, having waited for the present case, where an especially strong test of the point is to be found. In the present data, run is followed (within 2 sec. or less) by drink an average of about 50 times per session. Run is not followed by drink—being followed

either by another burst of run or by neither run nor drink for a period greater than 8 sec.—only about 6 times per 15-min. session. Moreover, 6 times per 15-min. session is about the same frequency with which this kind of event, an “unprotected” run in Guthrie’s sense, occurs in contingencies that are effective, that do reduce drink and do produce an increase in run. It is therefore difficult to argue that the effectiveness of the temporal contiguities between run and drink is dispelled in the present case by extinction from runs that are unaccompanied by drinks. On the contrary, run is followed repeatedly by drink—it is rarely followed by any other event—nevertheless, when drink is not reduced, run is not increased, and run shows no increment in subsequent would-be extinction sessions.

In summary, a set of operations were shown that lead to temporal contiguity between events that are incontrovertibly reinforcing and reforcible. If all such operations produced reinforcement, temporal contiguity could be accepted as a sufficient condition. As it turns out, however, only some of the operations produce reinforcement; we are thus led to search among the operations to discern the further conditions they embody. The present search indicates that a reduction in the total amount of H is a necessary condition. A response requirement, which is an essential part of the contingency, proved to be neither a necessary nor a sufficient condition. For example, if H is reduced, reinforcement can be produced without a response requirement, as in Skinner’s superstition paradigm. Conversely, if H is not reduced, then, despite the presence of both a response requirement and a circumscription on the distribution of H, reinforcement apparently cannot be produced, as shown by the special run-drink contingency where drink occurred in normal amount. Without doubt the contingency is the most widely used means of producing reinforcement. Apparently, an invariant though unrecognized component of the contingency is a decrement in the amount of responding that occurs to the contingent stimulus, relative to what would occur were the stimulus free. Our results suggest that this reduction is vital, that reinforcement cannot be initiated without it. We are thus led to suppose

that although the reduction is not, like the response requirement, a necessary feature of the contingency, it nonetheless occurs as a routine part of the reinforcement procedure. It would be preferable to elevate this factor from its obscure status as a hidden concomitant to that of a public operation where its consequences for theory can be examined.¹²

SUMMARY: COUNTEREXAMPLES

The argument has been too long to recapitulate easily, so let us summarize instead by briefly considering outstanding counterexamples.

Classical Conditioning

That the reinforcement value of a chain of operant behavior (or of the initiating discriminative stimulus) is proportional to the strength of the terminal respondent is a view that claims surprisingly many adherents, surprising given the slight evidence the view so far commands (for adherents, see Spence, 1956; Mowrer, 1960). Is this model compatible with the present one? Let us suspend that question to look first at the basic assumption of the model. Do operant chains terminate in reflexes? It is not necessary that all operant chains terminate in reflexes, but the view does require that those chains which participate in reinforcement do so terminate. In other words, the adequacy of the view requires that its domain be at least as broad as that of reinforcement. We can comment on that adequacy by looking at several of the operant chains which were shown here to participate in reinforcement.

Thus, we may ask what is the terminal respondent in manipulation of monkeys. Is it possible, in the case of the lever-pressing

¹² A sequel has shown that although such "weak" operations as pairing stimuli that control responses of different probabilities do not produce changes in frequency, they may nonetheless affect distributional properties; e.g., they may produce a quasi-permanent change in average duration of the burst. However, we have considered it to be less confusing to restrict the meaning of "reinforcement" here to the traditional frequency change. Elsewhere we shall deal with reinforcement under various relaxations of the standard definition.

or plunger operation of the present monkey experiment, to delineate the reflexes in which these undeniably operant chains terminated? What is the terminal respondent in wheel turning in rats? The impression grows that these questions, and more like them, may not be easily answered, especially when it is observed that a proper answer will require a description of the unconditioned stimulus, the location on the animal's body where it is to be applied, and the form of the response that is to result. Once reinforcement is loosened from eating and drinking—a loosening unlikely in those cases where the tie is essentially umbilical—we encounter cases where there is no visible respondent, no unconditioned stimulus to apply, no unconditioned response to measure. The adequacy of the view would thus seem to be confined mainly to the ingestive cases that spawned the view. But a theory of ingestion cannot be equated with a general model of reinforcement.

There is a more important point. The facts of classical conditioning are themselves in need of explanation. Rather than attempting to explain operant reinforcement with classical conditioning, we might instead look to the possibility of accounting for classical conditioning. It is at this point that an interesting possibility arises, viz., accounting for both classical conditioning and operant reinforcement by the same principle.

Standard treatments of operant reinforcement fail to make contact with the strong question "Why does (say) food reinforce the bar press?" being sidetracked instead by the weak question "Why is a large piece of food more effective than a small one?" But if one can answer the question of why X will reinforce Y, it should not be too difficult to explain why a large X may be more effective than a small one. An answer to the strong question is far more likely to include an answer to the weak question than vice versa.

Likewise, in classical conditioning the strong question tends to be overlooked in favor of the parameters of conditioning. Why are the responses elicited by some stimuli conditionable to other stimuli and not vice versa? That is, what makes some stimuli (utilizable as) unconditioned stimuli and others (utiliz-

able as) conditioned stimuli? The latter has been rather neglected, though Pavlov (1927) would appear to have regarded the question as basic and, along with Sherrington (1906), attempted to answer it. In fact, these men gave similar answers, both stressing the dominance of one stimulation relative to that of the other. More recent versions of dominance theory have been provided by Razran (1957) and Grings (1960).

We may now discern an evident similarity between the Pavlovian account of classical conditioning and the present one of operant reinforcement. A process is considered to depend upon a relationship between two events—US and CS in one case, base and contingent response in the other—and to require that one event be dominant relative to the other. It will be observed, however, that in order to complete the formal parallel, we must specify the dimension on which dominance is to be measured in the classical case. Pavlov does not appear to have made any operational provisions, having considered dominance mainly in neurological terms. There are only two operational alternatives: stimulus or response properties. If stimulus properties are used, dominance will be restricted to a few (though possibly paradigmatic) cases where US and CS can be located on the same scale. This is a restriction we specifically sought to overcome in the operant case by the use of response probability. But can the classical case be formulated in other than stimulus terms?

If analogues of the two responses in the operant case can be found in the classical one, the classical case too could be stated in response terms. There is no problem with the UR, for it is quite explicit and in fact sometimes measured; rather, the problem is with the response to the CS—not the CR, but the unconditional response to the CS. Is there such a response? Typical CSs such as lights and buzzers elicit reactions, of a lesser magnitude than those elicited by the typical US, which characteristically are ignored but which might instead be measured. The appropriate scale for their measurement has already been established, should measurement prove worthwhile, for the response rationale worked out for the operant case can be applied here without revision. That is, in the classical case, too, one will

encounter a set of physically different movements, all of which, however, will have duration or temporal extent in common. Thus, in principle at least, the decision as to dominance can be made on an equivalent basis in both cases.¹³

"Yes" as a Reinforcer

A case which questions the exhaustiveness more than the validity of the present account arises from a category of stimuli that are effective through another agent applying them to the subject rather than through the usual self-application. A simple example is the experimenter saying "yes" (or, for that matter, "wrong"; Buchwald, 1959) to a human subject. "Yes" does not appear to give rise to any determinate response on the part of the subject, and moreover it is not clear how to estimate the probability of *S*'s responding to *E*'s "yes." If we were to provide the subject with a "yes" machine—a box emitting "yes" whenever touched—*S* would seem unlikely to make extensive use of it, in particular, a use proportional to the reinforcement value of "yes" as applied by the experimenter. But if so, this would appear to refute the present account, not merely question its exhaustiveness. On the other hand, all that may be at stake here is a secondary reinforcer whose effectiveness depends upon a context more complex than that found in the usual animal example. A light made effective by association with a food pellet should display its effectiveness in any situation which sufficiently approximates the training one. Similarly, "yes," having acquired its potency in a distinctive linguistic relation, could be expected to reveal its effectiveness only in contexts that approximate the original linguistic relation. A "yes" box would not sufficiently approximate that relation. Furthermore, we would trace the

¹³ One immediate prediction which follows from thus equating the classical and operant cases is that the US and CS will no longer be seen as categories with fixed membership; on the contrary, if appropriate parameters can be found, it should be possible to reverse the classical conditioning relation in a manner comparable to the reversal of the reinforcement relation shown earlier in the operant case.

original effectiveness of "yes" to association with high-probability events.¹⁴

Nonresponse Reinforcement

A more blunt objection is found in data showing that, for example, fistula feeding is reinforcing (e.g., Miller & Kessen, 1952). Here the R-R paradigm is visibly no longer applicable: reinforcement occurs without a response. Such cases cannot be gainsaid, and we must instead take into account their one irrefutable implication. Organisms are of such a design that their responses can be circumvented. Indeed, the circumvention of all responses would seem only to await sufficient knowledge. In due time, the effects of, say, running will doubtless be duplicable without running, by neurological intervention, and these effects no less than those of eating or of any other behavior. To a non-specialist such as myself, the wholesale extent to which this intervention may be possible was nowhere better suggested than on a Sunday television program that showed a robot learn a maze. Having first struck nearly every blind, but finally learning to sail down smoothly, the little device was picked up and its head promptly opened. There were the dread insides—though in this case the black box *was* nearly empty—perfectly open to direct intervention. Never was it more clear that behavior is expendable. Although behavior is the locus of morality (that without which Conrad would have had no subject matter), it is nonetheless susceptible of total circumvention, contingent only upon sufficient knowledge. Strictly speaking, therefore, one cannot say, as we have, that a response is a necessary condition for

¹⁴ L. Homme (personal communication, 1965) has described an ingenious set of studies (using both normal children and mental defectives) in which speech was reinforced, not with the standard approval or attention, e.g., "yes," "umhum," but by making the opportunity to talk about a topic for which the probability of speech was high contingent upon talking about another topic for which the probability of speech was relatively low. For example, the opportunity to talk about baseball was made contingent upon the solution of arithmetic problems.

anything; in principle, one can always reach inside and pull a string.

Still one can ask whether it is possible to predict under what circumstance fistula feeding will be reinforcing. The present assumptions may regain their relevance at this point, for, although responses can indeed be circumvented, when they are not, measurements which they permit may yield generalizations that hold about as well for the surgical puppet as for the intact organism. Fistula feeding may prove to reinforce, say, bar pressing only if the probability of normal eating is greater than that of bar pressing. Alternatively, an organism that is less apt to eat than, say, scrutinize the passage of pinballs may be increased in the frequency with which it feeds itself through a fistula if the pinball opportunity is contingent upon self-fistula feeding. The response is not a necessary condition so much as an opportunity for measurement, one that would be regrettable to neglect for the information it may provide and the theory to which it may lead.

REFERENCES

- Bower, G. H., & Miller, N. E. Rewarding and punishing effects from stimulating the same place in a rat's brain. *J. comp. physiol. Psychol.*, 1958, **51**, 669-674.
- Buchwald, A. M. Experimental alternations in the effectiveness of reinforcement combinations. *J. exp. Psychol.*, 1959, **57**, 351-361.
- Catania, A. C. Concurrent performances: reinforcement interaction and response independence. *J. exp. Anal. Behav.*, 1963, **6**, 253-264. (a)
- Catania, A. C. Concurrent performances: a baseline for the study of reinforcement magnitude. *J. exp. Anal. Behav.*, 1963, **6**, 299-301. (b)
- Collier, G. Consummatory and instrumental responding as a function of deprivation. *J. exp. Psychol.*, 1962, **64**, 410-414.
- Collier, G., & Siskel, M., Jr. Performance as a joint function of amount of reinforcement and inter-reinforcement interval. *J. exp. Psychol.*, 1959, **57**, 115-120.
- Dethier, V. G. Microscopic brains. *Science*, 1964, **143**, 1138-1145.
- Grings, W. W. Preparatory set variables in the classical conditioning of autonomic variables. *Psychol. Rev.*, 1960, **67**, 243-252.
- Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. exp. Anal. Behav.*, 1961, **4**, 267-273.
- Hodos, W., & Valenstein, E. S. An evaluation of response rate as a measure of rewarding intracranial stimulation. *J. comp. physiol. Psychol.*, 1962, **55**, 80-84.

- Holstein, S., & Hundt, A. G. Reinforcement of intracranial self-stimulation by licking. *Psychonomic Sc.*, in press.
- Hundt, A. G. Instrumental response rate and reinforcement density. Unpublished doctoral dissertation, Univ. of Missouri, 1964.
- Hundt, A. G., & Premack, D. Running as both a positive and negative reinforcer. *Science*, 1963, 142, 1087-1088.
- Kintsch, W., & Premack, D. Stochastic analysis of free responding. Paper read at Psychometric Society, Niagara Falls, 1965.
- Meehl, P. E. On the circularity of the law of effect. *Psychol. Bull.*, 1950, 47, 52-75.
- Meehl, P. E. Schizotaxia, Schizotypy, Schizophrenia. *Amer. J. Psychol.*, 1962, 17, 827-838.
- Miller, N., & Kessen, M. L. Reward effects of food via stomach fistula compared with those of food via mouth. *J. comp. physiol. Psychol.*, 1952, 45, 555-564.
- Mowrer, O. H. *Learning theory and behavior*. New York: Wiley, 1960.
- Parzen, E. *Modern probability theory and its applications*. New York: Wiley, 1960.
- Pavlov, I. P. *Conditioned reflexes*. (Trans. by G. V. Anrep) London: Oxford Univ. Press, 1927.
- Premack, D. Toward empirical behavioral laws: I. Positive reinforcement. *Psychol. Rev.*, 1959, 66, 219-233.
- Premack, D. Predicting instrumental performance from the independent rate of the contingent response. *J. exp. Psychol.*, 1961, 61, 163-171.
- Premack, D. Reversibility of the reinforcement relation. *Science*, 1962, 136, 255-257.
- Premack, D. Rate differential reinforcement in monkey manipulation. *J. exp. Anal. Behav.*, 1963, 6, 81-89. (a)
- Premack, D. Prediction of the comparative reinforcement values of running and drinking. *Science*, 1963, 139, 1062-1063. (b)
- Premack, D., & Bahwell, R. Operant-level lever pressing by a monkey as a function of interest interval. *J. exp. Anal. Behav.*, 1959, 2, 127-131.
- Premack, D., & Schaeffer, R. W. Distributional properties of operant-level locomotion in the rat. *J. exp. Anal. Behav.*, 1962, 5, 89-95.
- Premack, D., & Schaeffer, R. W. Some parameters affecting the distributional properties of operant-level running in rats. *J. exp. Anal. Behav.*, 1963, 6, 473-475.
- Pribram, K. H. Reinforcement revisited: a structural view. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1963*. Lincoln: Univ. of Nebraska Press, 1963. Pp. 113-150.
- Razran, G. The dominance-contiguity theory of the acquisition of classical conditioning. *Psychol. Bull.*, 1957, 54, 1-46.
- Reynolds, G. S. On some determinants of choice in pigeons. *J. exp. Anal. Behav.*, 1963, 6, 53-62.
- Roberts, W. W. Rapid escape learning without avoidance learning motivated

- by hypothalamic stimulation in cats. *J. comp. physiol. Psychol.*, 1958, 51, 391-399.
- Schaeffer, R. W. The reinforcement relation as a function of the instrumental response base rate. *J. exp. Psychol.*, in press.
- Schaeffer, R. W., & Premack, D. Licking rates in infant albino rats. *Science*, 1961, 134, 1980-1981.
- Sherrington, C. S. *The integrative action of the nervous system*. New Haven: Yale Univ. Press, 1906.
- Skinner, B. F. The measurement of spontaneous activity. *J. gen. Psychol.*, 1933, 9, 3-24.
- Skinner, B. F. *The behavior of organisms*. New York: Appleton-Century-Crofts, 1938.
- Skinner, B. F. "Superstition" in the pigeon. *J. exp. Psychol.*, 1948, 38, 168-172.
- Spence, K. W. *Behavior theory and conditioning*. New Haven: Yale Univ. Press, 1956.
- Stellar, E., & Hill, J. H. The rat's rate of drinking as a function of water deprivation. *J. comp. physiol. Psychol.*, 1952, 45, 96-102.

Comments on Dr. Premack's Paper

Robert Ward Leeper

There are a few questions I want to raise about Dr. Premack's paper, but let me say, first, that I regard this as an unusually keen and interesting paper—one that makes some enduring contributions to the psychology of motivation. Let me speak first about the features that prompt this remark.

First of all, I must confess that I am highly sympathetic with certain temperamental characteristics that run through Premack's paper. I like his readiness—his eagerness, even—to challenge traditional conceptions in psychology and to seek better alternatives. I admire his predilection for casting his hypotheses in very broad, abstract terms rather than timidly confining his statements to the particular conditions of his particular experiments. I would have preferred for him to state these abstract propositions in simpler, plainer terms, and without the shifting meanings he gives to "reinforcement" and to his "H" and "L." But these matters are

Premack, D. (1965). Reinforcement theory. In D. Levine (Ed.), *Nebraska symposium on motivation* (vol. 13, pp. 123-180). Lincoln: University of Nebraska.