

THE PSYCHOLOGICAL REVIEW

ARE THEORIES OF LEARNING NECESSARY? ¹

BY B. F. SKINNER

Harvard University

Certain basic assumptions, essential to any scientific activity, are sometimes called theories. That nature is orderly rather than capricious is an example. Certain statements are also theories simply to the extent that they are not yet facts. A scientist may guess at the result of an experiment before the experiment is carried out. The prediction and the later statement of result may be composed of the same terms in the same syntactic arrangement, the difference being in the degree of confidence. No empirical statement is wholly non-theoretical in this sense, because evidence is never complete, nor is any prediction probably ever made wholly without evidence. The term "theory" will not refer here to statements of these sorts but rather to any explanation of an observed fact which appeals to events taking place somewhere else, at some other level of observation, described in different terms, and measured, if at all, in different dimensions.

Three types of theory in the field of learning satisfy this definition. The most characteristic is to be found in the field of physiological psychology. We are all familiar with the changes that are supposed to take place in the nervous system when an organism learns.

Synaptic connections are made or broken, electrical fields are disrupted or reorganized, concentrations of ions are built up or allowed to diffuse away, and so on. In the science of neurophysiology statements of this sort are not necessarily theories in the present sense. But in a science of behavior, where we are concerned with whether or not an organism secretes saliva when a bell rings, or jumps toward a gray triangle, or says *bik* when a cards reads *tuz*, or loves someone who resembles his mother, all statements about the nervous system are theories in the sense that they are not expressed in the same terms and could not be confirmed with the same methods of observation as the facts for which they are said to account.

A second type of learning theory is in practice not far from the physiological, although there is less agreement about the method of direct observation. Theories of this type have always dominated the field of human behavior. They consist of references to "mental" events, as in saying that an organism learns to behave in a certain way because it "finds something pleasant" or because it "expects something to happen." To the mentalistic psychologist these explanatory events are no more theoretical than synaptic connections to the neurophysiologist, but in a science of behavior they are theories because the methods

¹ Address of the president, Midwestern Psychological Association, Chicago, Illinois, May, 1949.

and terms appropriate to the events to be explained differ from the methods and terms appropriate to the explaining events.

In a third type of learning theory the explanatory events are not directly observed. The writer's suggestion that the letters CNS be regarded as representing, not the Central Nervous System, but the Conceptual Nervous System (2, p. 421), seems to have been taken seriously. Many theorists point out that they are not talking about the nervous system as an actual structure undergoing physiological or bio-chemical changes but only as a system with a certain dynamic output. Theories of this sort are multiplying fast, and so are parallel operational versions of mental events. A purely behavioral definition of expectancy has the advantage that the problem of mental observation is avoided and with it the problem of how a mental event can cause a physical one. But such theories do not go so far as to assert that the explanatory events are identical with the behavioral facts which they purport to explain. A statement about behavior may support such a theory but will never resemble it in terms or syntax. Postulates are good examples. True postulates cannot become facts. Theorems may be deduced from them which, as tentative statements about behavior, may or may not be confirmed, but theorems are not theories in the present sense. Postulates remain theories until the end.

It is not the purpose of this paper to show that any of these theories cannot be put in good scientific order, or that the events to which they refer may not actually occur or be studied by appropriate sciences. It would be foolhardy to deny the achievements of theories of this sort in the history of science. The question of whether they are necessary, however, has other implications and is worth asking. If the answer is no, then

it may be possible to argue effectively against theory in the field of learning. A science of behavior must eventually deal with behavior in its relation to certain manipulable variables. Theories—whether neural, mental, or conceptual—talk about intervening steps in these relationships. But instead of prompting us to search for and explore relevant variables, they frequently have quite the opposite effect. When we attribute behavior to a neural or mental event, real or conceptual, we are likely to forget that we still have the task of accounting for the neural or mental event. When we assert that an animal acts in a given way because it expects to receive food, then what began as the task of accounting for learned behavior becomes the task of accounting for expectancy. The problem is at least equally complex and probably more difficult. We are likely to close our eyes to it and to use the theory to give us answers in place of the answers we might find through further study. It might be argued that the principal function of learning theory to date has been, not to suggest appropriate research, but to create a false sense of security, an unwarranted satisfaction with the *status quo*.

Research designed with respect to theory is also likely to be wasteful. That a theory generates research does not prove its value unless the research is valuable. Much useless experimentation results from theories, and much energy and skill are absorbed by them. Most theories are eventually overthrown, and the greater part of the associated research is discarded. This could be justified if it were true that productive research requires a theory, as is, of course, often claimed. It is argued that research would be aimless and disorganized without a theory to guide it. The view is supported by psychological texts that take their cue from the logicians rather than empirical science and

describe thinking as necessarily involving stages of hypothesis, deduction, experimental test, and confirmation. But this is not the way most scientists actually work. It is possible to design significant experiments for other reasons and the possibility to be examined is that such research will lead more directly to the kind of information that a science usually accumulates.

The alternatives are at least worth considering. How much can be done without theory? What other sorts of scientific activity are possible? And what light do alternative practices throw upon our present preoccupation with theory?

It would be inconsistent to try to answer these questions at a theoretical level. Let us therefore turn to some experimental material in three areas in which theories of learning now flourish and raise the question of the function of theory in a more concrete fashion.²

The Basic Datum in Learning

What actually happens when an organism learns is not an easy question. Those who are interested in a science of behavior will insist that learning is a change in behavior, but they tend to avoid explicit references to responses or acts as such. "Learning is adjustment, or adaptation to a situation." But of what stuff are adjustments and adaptations made? Are they data, or inferences from data? "Learning is improvement." But improvement in what? And from whose point of view? "Learning is restoration of equilibrium." But what

² Some of the material that follows was obtained in 1941-42 in a cooperative study on the behavior of the pigeon in which Keller Breland, Norman Guttman, and W. K. Estes collaborated. Some of it is selected from subsequent, as yet unpublished, work on the pigeon conducted by the author at Indiana University and Harvard University. Limitations of space make it impossible to report full details here.

is in equilibrium and how is it put there? "Learning is problem solving." But what are the physical dimensions of a problem—or of a solution? Definitions of this sort show an unwillingness to take what appears before the eyes in a learning experiment as a basic datum. Particular observations seem too trivial. An error score falls; but we are not ready to say that this is learning rather than merely the result of learning. An organism meets a criterion of ten successful trials; but an arbitrary criterion is at variance with our conception of the generality of the learning process.

This is where theory steps in. If it is not the time required to get out of a puzzle-box that changes in learning, but rather the strength of a bond, or the conductivity of a neural pathway, or the excitatory potential of a habit, then problems seem to vanish. Getting out of a box faster and faster is not learning; it is merely performance. The learning goes on somewhere else, in a different dimensional system. And although the time required depends upon arbitrary conditions, often varies discontinuously, and is subject to reversals of magnitude, we feel sure that the learning process itself is continuous, orderly, and beyond the accidents of measurement. Nothing could better illustrate the use of theory as a refuge from the data.

But we must eventually get back to an observable datum. If learning is the process we suppose it to be, then it must appear so in the situations in which we study it. Even if the basic process belongs to some other dimensional system, our measures must have relevant and comparable properties. But productive experimental situations are hard to find, particularly if we accept certain plausible restrictions. To show an orderly change in the behavior of the *average* rat or ape or child is not enough, since learning is a process in the behavior of

the individual. To record the beginning and end of learning or a few discrete steps will not suffice, since a series of cross-sections will not give complete coverage of a continuous process. The dimensions of the change must spring from the behavior itself; they must not be imposed by an external judgment of success or failure or an external criterion of completeness. But when we review the literature with these requirements in mind, we find little justification for the theoretical process in which we take so much comfort.

The energy level or work-output of behavior, for example, does not change in appropriate ways. In the sort of behavior adapted to the Pavlovian experiment (respondent behavior) there may be a progressive increase in the magnitude of response during learning. But we do not shout our responses louder and louder as we learn verbal material, nor does a rat press a lever harder and harder as conditioning proceeds. In operant behavior the energy or magnitude of response changes significantly only when some arbitrary value is differentially reinforced—when such a change is what is learned.

The emergence of a right response in competition with wrong responses is another datum frequently used in the study of learning. The maze and the discrimination box yield results which may be reduced to these terms. But a behavior-ratio of right *vs.* wrong cannot yield a continuously changing measure in a single experiment on a single organism. The point at which one response takes precedence over another cannot give us the whole history of the change in either response. Averaging curves for groups of trials or organisms will not solve this problem.

Increasing attention has recently been given to latency, the relevance of which, like that of energy level, is suggested by the properties of conditioned and uncon-

ditioned reflexes. But in operant behavior the relation to a stimulus is different. A measure of latency involves other considerations, as inspection of any case will show. Most operant responses may be emitted in the absence of what is regarded as a relevant stimulus. In such a case the response is likely to appear before the stimulus is presented. It is no solution to escape this embarrassment by locking a lever so that an organism cannot press it until the stimulus is presented, since we can scarcely be content with temporal relations that have been forced into compliance with our expectations. Runway latencies are subject to this objection. In a typical experiment the door of a starting box is opened and the time that elapses before a rat leaves the box is measured. Opening the door is not only a stimulus, it is a change in the situation that makes the response possible for the first time. The time measured is by no means as simple as a latency and requires another formulation. A great deal depends upon what the rat is doing at the moment the stimulus is presented. Some experimenters wait until the rat is facing the door, but to do so is to tamper with the measurement being taken. If, on the other hand, the door is opened without reference to what the rat is doing, the first major effect is the conditioning of favorable waiting behavior. The rat eventually stays near and facing the door. The resulting shorter starting-time is not due to a reduction in the latency of a response, but to the conditioning of favorable preliminary behavior.

Latencies in a single organism do not follow a simple learning process. Relevant data on this point were obtained as part of an extensive study of reaction time. A pigeon, enclosed in a box, is conditioned to peck at a recessed disc in one wall. Food is presented as reinforcement by exposing a hopper through

a hole below the disc. If responses are reinforced only after a stimulus has been presented, responses at other times disappear. Very short reaction times are obtained by differentially reinforcing responses which occur very soon after the stimulus (4). But responses also come to be made very quickly without differential reinforcement. Inspection shows that this is due to the development of effective waiting. The bird comes to stand before the disc with its head in good striking position. Under optimal conditions, without differential reinforcement, the mean time between stimulus and response will be of the order of $\frac{1}{3}$ sec. This is not a true reflex latency, since the stimulus is discriminative rather than eliciting, but it is a fair example of the latency used in the study of learning. The point is that this measure does not vary continuously or in an orderly fashion. By giving the bird more food, for example, we induce a condition in which it does not always respond. But the responses that occur show approximately the same temporal relation to the stimulus (Fig. 1, middle curve). In extinction, of special interest here, there is a scattering of latencies because lack of reinforcement generates an emotional condition.

Some responses occur sooner and others are delayed, but the commonest value remains unchanged (bottom curve in Fig. 1). The longer latencies are easily explained by inspection. Emotional behavior, of which examples will be mentioned later, is likely to be in progress when the ready-signal is presented. It is often not discontinued before the "go" signal is presented, and the result is a long starting-time. Cases also begin to appear in which the bird simply does not respond at all during a specified time. If we average a large number of readings, either from one bird or many, we may create what looks like a progressive lengthening of latency. But the data for an individual organism do not show a continuous process.

Another datum to be examined is the rate at which a response is emitted. Fortunately the story here is different. We study this rate by designing a situation in which a response may be freely repeated, choosing a response (for example, touching or pressing a small lever or key) that may be easily observed and counted. The responses may be recorded on a polygraph, but a more convenient form is a cumulative curve from which rate of responding is immediately read as slope. The rate at which a response is emitted in such a situation comes close to our preconception of the learning process. As the organism learns, the rate rises. As it unlearns (for example, in extinction) the rate falls. Various sorts of discriminative stimuli may be brought into control of the response with corresponding modifications of the rate. Motivational changes alter the rate in a sensitive way. So do those events which we speak of as generating emotion. The range through which the rate varies significantly may be as great as of the order of 1000:1. Changes in rate are satisfactorily smooth in the individual case, so that it is not necessary to aver-

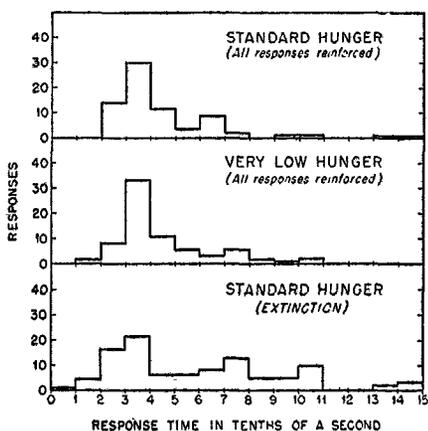


FIG. 1

age cases. A given value is often quite stable: in the pigeon a rate of four or five thousand responses per hour may be maintained without interruption for as long as fifteen hours.

Rate of responding appears to be the only datum that varies significantly and in the expected direction under conditions which are relevant to the "learning process." We may, therefore, be tempted to accept it as our long-sought-for measure of strength of bond, excitatory potential, etc. Once in possession of an effective datum, however, we may feel little need for any theoretical construct of this sort. Progress in a scientific field usually waits upon the discovery of a satisfactory dependent variable. Until such a variable has been discovered, we resort to theory. The entities which have figured so prominently in learning theory have served mainly as substitutes for a directly observable and productive datum. They have little reason to survive when such a datum has been found.

It is no accident that rate of responding is successful as a datum, because it is particularly appropriate to the fundamental task of a science of behavior. If we are to predict behavior (and possibly to control it), we must deal with *probability of response*. The business of a science of behavior is to evaluate this probability and explore the conditions that determine it. Strength of bond, expectancy, excitatory potential, and so on, carry the notion of probability in an easily imagined form, but the additional properties suggested by these terms have hindered the search for suitable measures. Rate of responding is not a "measure" of probability but it is the only appropriate datum in a formulation in these terms.

As other scientific disciplines can attest, probabilities are not easy to handle. We wish to make statements about the likelihood of occurrence of a single

future response, but our data are in the form of frequencies of responses that have already occurred. These responses were presumably similar to each other and to the response to be predicted. But this raises the troublesome problem of response-instance *vs.* response-class. Precisely what responses are we to take into account in predicting a future instance? Certainly not the responses made by a population of different organisms, for such a statistical datum raises more problems than it solves. To consider the frequency of repeated responses in an individual demands something like the experimental situation just described.

This solution of the problem of a basic datum is based upon the view that operant behavior is essentially an emissive phenomenon. Latency and magnitude of response fail as measures because they do not take this into account. They are concepts appropriate to the field of the reflex, where the all but invariable control exercised by the eliciting stimulus makes the notion of probability of response trivial. Consider, for example, the case of latency. Because of our acquaintance with simple reflexes we infer that a response that is more likely to be emitted will be emitted more quickly. But is this true? What can the word "quickly" mean? Probability of response, as well as prediction of response, is concerned with the moment of emission. This is a point in time, but it does not have the temporal dimension of a latency. The execution may take time after the response has been initiated, but the moment of occurrence has no duration.³ In recog-

³ It cannot, in fact, be shortened or lengthened. Where a latency appears to be forced toward a minimal value by differential reinforcement, another interpretation is called for. Although we may differentially reinforce more energetic behavior or the faster execution of behavior after it begins, it is meaningless to speak of differentially reinforcing re-

nizing the emissive character of operant behavior and the central position of probability of response as a datum, latency is seen to be irrelevant to our present task.

Various objections have been made to the use of rate of responding as a basic datum. For example, such a program may seem to bar us from dealing with many events which are unique occurrences in the life of the individual. A man does not decide upon a career, get married, make a million dollars, or get killed in an accident often enough to make a rate of response meaningful. But these activities are not responses. They are not simple unitary events lending themselves to prediction as such. If we are to predict marriage, success, accidents, and so on, in anything more than statistical terms, we must deal with the smaller units of behavior which lead to and compose these unitary episodes. If the units appear in repeatable form, the present analysis may be applied. In the field of learning a similar objection takes the form of asking how the present analysis may be extended to experimental situations in which it is impossible to observe frequencies. It does

sponses with short or long latencies. What we actually reinforce differentially are (a) favorable waiting behavior and (b) more vigorous responses. When we ask a subject to respond "as soon as possible" in the human reaction-time experiment, we essentially ask him (a) to carry out as much of the response as possible without actually reaching the criterion of emission, (b) to do as little else as possible, and (c) to respond energetically after the stimulus has been given. This may yield a minimal measurable time between stimulus and response, but this time is not necessarily a basic datum nor have our instructions altered it as such. A parallel interpretation of the differential reinforcement of long "latencies" is required. This is easily established by inspection. In the experiments with pigeons previously cited, preliminary behavior is conditioned that postpones the response to the key until the proper time. Behavior that "marks time" is usually conspicuous,

not follow that learning is not taking place in such situations. The notion of probability is usually extrapolated to cases in which a frequency analysis cannot be carried out. In the field of behavior we arrange a situation in which frequencies are available as data, but we use the notion of probability in analyzing and formulating instances or even types of behavior which are not susceptible to this analysis.

Another common objection is that a rate of response is just a set of latencies and hence not a new datum at all. This is easily shown to be wrong. When we measure the time elapsing between two responses, we are in no doubt as to what the organism was doing when we started our clock. We know that it was just executing a response. This is a natural zero—quite unlike the arbitrary point from which latencies are measured. The free repetition of a response yields a rhythmic or periodic datum very different from latency. Many periodic physical processes suggest parallels.

We do not choose rate of responding as a basic datum merely from an analysis of the fundamental task of a science of behavior. The ultimate appeal is to its success in an experimental science. The material which follows is offered as a sample of what can be done. It is not intended as a complete demonstration, but it should confirm the fact that when we are in possession of a datum which varies in a significant fashion, we are less likely to resort to theoretical entities carrying the notion of probability of response.

Why Learning Occurs

We may define learning as a change in probability of response but we must also specify the conditions under which it comes about. To do this we must survey some of the independent variables of which probability of response is

a function. Here we meet another kind of learning theory.

An effective class-room demonstration of the Law of Effect may be arranged in the following way. A pigeon, reduced to 80 per cent of its *ad lib* weight, is habituated to a small, semi-circular amphitheatre and is fed there for several days from a food hopper, which the experimenter presents by closing a hand switch. The demonstration consists of establishing a selected response by suitable reinforcement with food. For example, by sighting across the amphitheatre at a scale on the opposite wall, it is possible to present the hopper whenever the top of the pigeon's head rises above a given mark. Higher and higher marks are chosen until, within a few minutes, the pigeon is walking about the cage with its head held as high as possible. In another demonstration the bird is conditioned to strike a marble placed on the floor of the amphitheatre. This may be done in a few minutes by reinforcing successive steps. Food is presented first when the bird is merely moving near the marble, later when it looks down in the direction of the marble, later still when it moves its head toward the marble, and finally when it pecks it. Anyone who has seen such a demonstration knows that the Law of Effect is no theory. It simply specifies a procedure for altering the probability of a chosen response.

But when we try to say *why* reinforcement has this effect, theories arise. Learning is said to take place because the reinforcement is pleasant, satisfying, tension reducing, and so on. The converse process of extinction is explained with comparable theories. If the rate of responding is first raised to a high point by reinforcement and reinforcement then withheld, the response is observed to occur less and less frequently thereafter. One common theory explains this by asserting that a state is

built up which suppresses the behavior. This "experimental inhibition" or "reaction inhibition" must be assigned to a different dimensional system, since nothing at the level of behavior corresponds to opposed processes of excitation and inhibition. Rate of responding is simply increased by one operation and decreased by another. Certain effects commonly interpreted as showing release from a suppressing force may be interpreted in other ways. Disinhibition, for example, is not necessarily the uncovering of suppressed strength; it may be a sign of supplementary strength from an extraneous variable. The process of spontaneous recovery, often cited to support the notion of suppression, has an alternative explanation, to be noted in a moment.

Let us evaluate the question of why learning takes place by turning again to some data. Since conditioning is usually too rapid to be easily followed, the process of extinction will provide us with a more useful case. A number of different types of curves have been consistently obtained from rats and pigeons using various schedules of prior reinforcement. By considering some of the relevant conditions we may see what room is left for theoretical processes.

The mere passage of time between conditioning and extinction is a variable that has surprisingly little effect. The rat is too short-lived to make an extended experiment feasible, but the pigeon, which may live ten or fifteen years, is an ideal subject. More than five years ago, twenty pigeons were conditioned to strike a large translucent key upon which a complex visual pattern was projected. Reinforcement was contingent upon the maintenance of a high and steady rate of responding and upon striking a particular feature of the visual pattern. These birds were set aside in order to study retention. They were transferred to the usual living

quarters, where they served as breeders. Small groups were tested for extinction at the end of six months, one year, two years, and four years. Before the test each bird was transferred to a separate living cage. A controlled feeding schedule was used to reduce the weight to approximately 80 per cent of the *ad lib* weight. The bird was then fed in the dimly lighted experimental apparatus in the absence of the key for several days, during which emotional responses to the apparatus disappeared. On the day of the test the bird was placed in the darkened box. The translucent key was present but not lighted. No responses were made. When the pattern was projected upon the key, all four birds responded quickly and extensively. Fig. 2 shows the largest curve obtained. This bird struck the key within two seconds after presentation of a visual pattern that it had not seen for four years, and at the precise spot upon which differential reinforcement had previously been based. It continued to respond for the next hour, emitting about 700 responses. This is of the order of one-half to one-quarter of the responses it would have emitted if extinction had not been delayed four years, but otherwise, the curve is fairly typical.

Level of motivation is another variable to be taken into account. An example of the effect of hunger has been

reported elsewhere (3). The response of pressing a lever was established in eight rats with a schedule of periodic reinforcement. They were fed the main part of their ration on alternate days so that the rates of responding on successive days were alternately high and low. Two subgroups of four rats each were matched on the basis of the rate maintained under periodic reinforcement under these conditions. The response was then extinguished—in one group on alternate days when the hunger was high, in the other group on alternate days when the hunger was low. (The same amount of food was eaten on the non-experimental days as before.) The result is shown in Fig. 3. The upper graph gives the raw data. The levels of hunger are indicated by the points at P on the abscissa, the rates prevailing under periodic reinforcement. The subsequent points show the decline in extinction. If we multiply the lower curve through by a factor chosen to superimpose the points at P, the curves are reasonably closely superimposed, as shown in the lower graph. Several other experiments on both rats and pigeons have confirmed this general principle. If a given ratio of responding prevails under periodic reinforcement, the slopes of later extinction curves show the same ratio. Level of hunger determines the slope of the extinction curve but not its curvature.

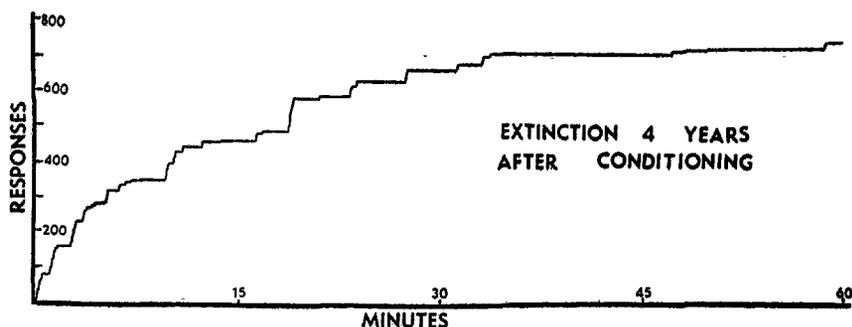


FIG. 2

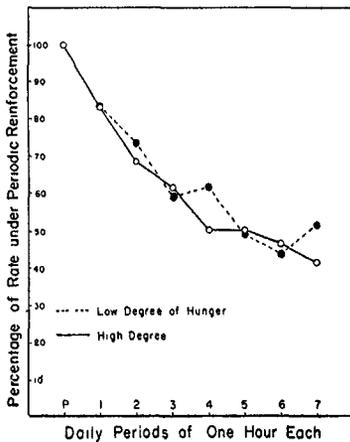
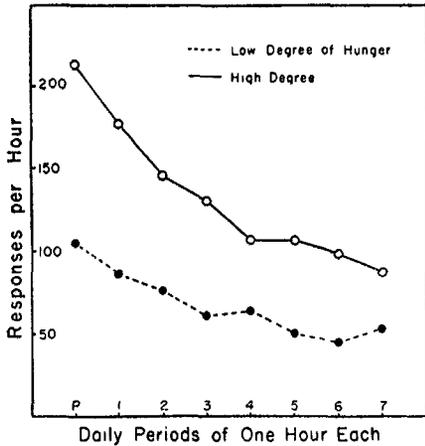


FIG. 3

Another variable, difficulty of response, is especially relevant because it has been used to test the theory of reaction inhibition (1), on the assumption that a response requiring considerable energy will build up more reaction inhibition than an easy response and lead, therefore, to faster extinction. The theory requires that the curvature of the extinction curve be altered, not merely its slope. Yet there is evidence that difficulty of response acts like level of hunger simply to alter the slope. Some data have been reported but not published (5). A pigeon is suspended in a jacket which confines its wings and

legs but leaves its head and neck free to respond to a key and a food magazine. Its behavior in this situation is quantitatively much like that of a bird moving freely in an experimental box. But the use of the jacket has the advantage that the response to the key may be made easy or difficult by changing the distance the bird must reach. In one experiment these distances were expressed in seven equal but arbitrary units. At distance 7 the bird could barely reach the key, at 3 it could strike without appreciably extending its neck. Periodic reinforcement gave a straight base-line upon which it was possible to observe the effect of difficulty by quickly changing position during the experimental period. Each of the five records in Fig. 4 covers a fifteen minute experimental period under periodic reinforcement. Distances of the bird from the key are indicated by numerals above the records. It will be observed that the rate of responding at distance 7 is generally quite low while that at distance 3 is high. Intermediate distances produce intermediate slopes. It should also be noted that the change from one position to another is felt immediately. If repeated responding in a difficult position were to build a considerable amount of reaction inhibition, we should expect the rate to be low for some little time after returning to an easy response. Contrariwise, if an easy response were to build little reaction inhibition, we should expect a fairly high rate of responding for some time after a difficult position is assumed. Nothing like this occurs. The "more rapid extinction" of a difficult response is an ambiguous expression. The slope constant is affected and with it the number of responses in extinction to a criterion, but there may be no effect upon curvature.

One way of considering the question of why extinction curves are curved is to regard extinction as a process of ex-

haustion comparable to the loss of heat from source to sink or the fall in the level of a reservoir when an outlet is opened. Conditioning builds up a predisposition to respond—a “reserve”—which extinction exhausts. This is perhaps a defensible description at the level of behavior. The reserve is not necessarily a theory in the present sense, since it is not assigned to a different dimensional system. It could be operationally defined as a predicted extinction curve, even though, linguistically, it makes a statement about the momentary condition of a response. But it is not a

particularly useful concept, nor does the view that extinction is a process of exhaustion add much to the observed fact that extinction curves are curved in a certain way.

There are, however, two variables that affect the rate, both of which operate during extinction to alter the curvature. One of these falls within the field of emotion. When we fail to reinforce a response that has previously been reinforced, we not only initiate a process of extinction, we set up an emotional response—perhaps what is often meant by frustration. The pigeon coos in an

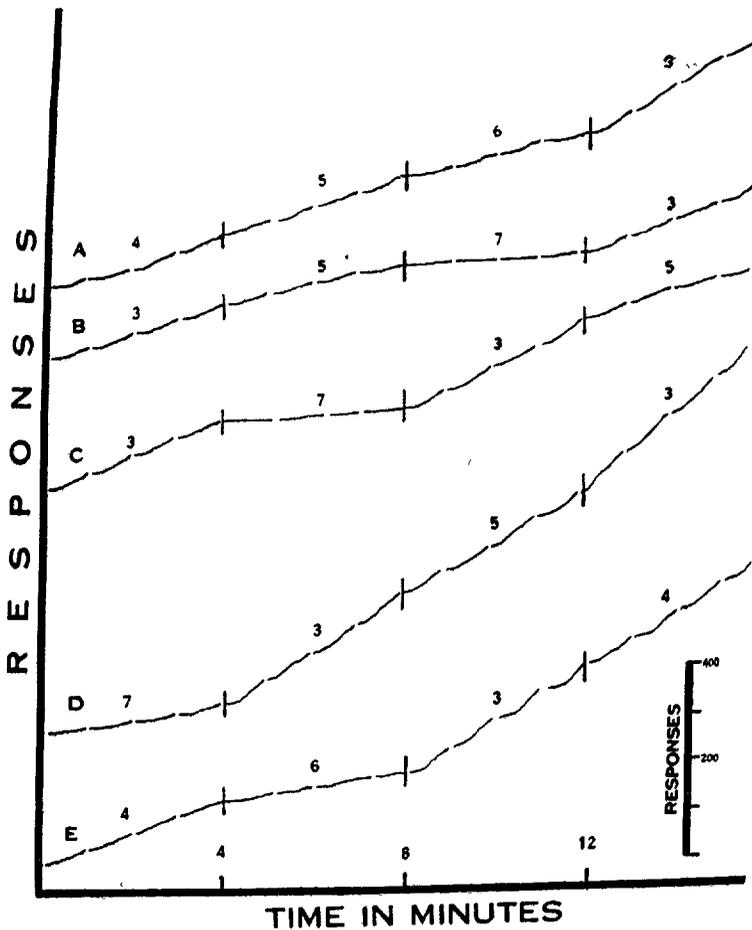


FIG. 4

identifiable pattern, moves rapidly about the cage, defecates, or flaps its wings rapidly in a squatting position that suggests treading (mating) behavior. This competes with the response of striking a key and is perhaps enough to account for the decline in rate in early extinction. It is also possible that the probability of a response based upon food deprivation is directly reduced as part of such an emotional reaction. Whatever its nature, the effect of this variable is eliminated through adaptation. Repeated extinction curves become smoother, and in some of the schedules to be described shortly there is little or no evidence of an emotional modification of rate.

A second variable has a much more serious effect. Maximal responding during extinction is obtained only when the conditions under which the response was reinforced are precisely reproduced. A rat conditioned in the presence of a light will not extinguish fully in the absence of the light. It will begin to respond more rapidly when the light is again introduced. This is true for other kinds of stimuli, as the following classroom experiment illustrates. Nine pigeons were conditioned to strike a yellow triangle under intermittent reinforcement. In the session represented by Fig. 5 the birds were first reinforced on this schedule for 30 minutes. The combined cumulative curve is essentially a straight line, showing more than 1100 responses per bird during this period. A red triangle was then substituted for the yellow and no responses were reinforced thereafter. The effect was a sharp drop in responding, with only a slight recovery during the next fifteen minutes. When the yellow triangle was replaced, rapid responding began immediately and the usual extinction curve followed. Similar experiments have shown that the pitch of an incidental tone, the shape of a pattern being

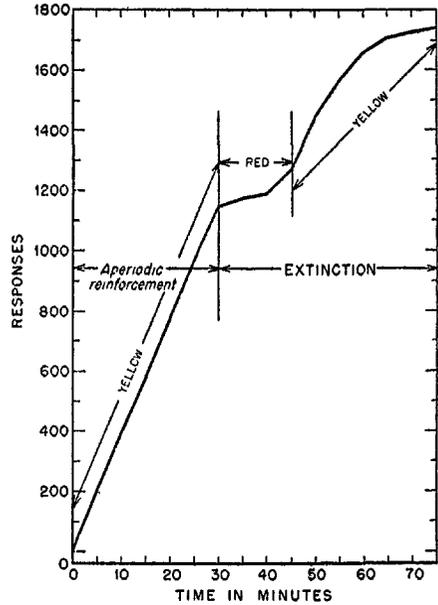


FIG. 5

struck, or the size of a pattern, if present during conditioning, will to some extent control the rate of responding during extinction. Some properties are more effective than others, and a quantitative evaluation is possible. By changing to several values of a stimulus in random order repeatedly during the extinction process, the gradient for stimulus generalization may be read directly in the rates of responding under each value.

Something very much like this must go on during extinction. Let us suppose that all responses to a key have been reinforced and that each has been followed by a short period of eating. When we extinguish the behavior, we create a situation in which responses are not reinforced, in which no eating takes place, and in which there are probably new emotional responses. The situation could easily be as novel as a red triangle after a yellow. If so, it could explain the decline in rate during extinction. We might have obtained a

smooth curve, *shaped like an extinction curve*, between the vertical lines in Fig. 5 by *gradually* changing the color of the triangle from yellow to red. This might have happened even though no other sort of extinction were taking place. The very conditions of extinction seem to presuppose a growing novelty in the experimental situation. Is this why the extinction curve is curved?

Some evidence comes from the data of "spontaneous recovery." Even after prolonged extinction an organism will often respond at a higher rate for at least a few moments at the beginning of another session. One theory contends that this shows spontaneous recovery from some sort of inhibition, but another explanation is possible. No matter how carefully an animal is handled, the stimulation coincident with the beginning of an experiment must be extensive and unlike anything occurring in the later part of an experimental period. Responses have been reinforced in the presence of, or shortly following, the organism is again placed in the experimental situation, the stimulation is this stimulation. In extinction it is present for only a few moments. When

restored; further responses are emitted as in the case of the yellow triangle. The only way to achieve full extinction in the presence of the stimulation of starting an experiment is to start the experiment repeatedly.

Other evidence of the effect of novelty comes from the study of periodic reinforcement. The fact that intermittent reinforcement produces bigger extinction curves than continuous reinforcement is a troublesome difficulty for those who expect a simple relation between number of reinforcements and number of responses in extinction. But this relation is actually quite complex. One result of periodic reinforcement is that emotional changes adapt out. This may be responsible for the smoothness of subsequent extinction curves but probably not for their greater extent. The latter may be attributed to the lack of novelty in the extinction situation. Under periodic reinforcement many responses are made without reinforcement and when no eating has recently taken place. The situation in extinction is therefore not wholly novel.

Periodic reinforcement is not, however, a simple solution. If we reinforce

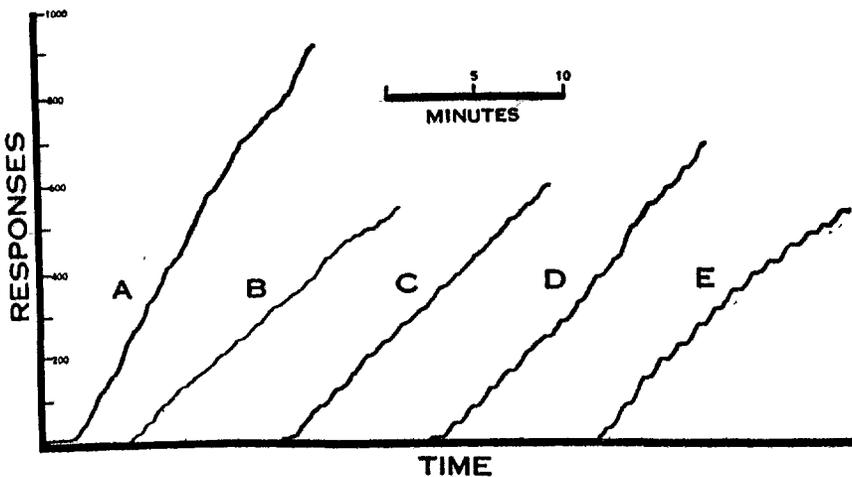


FIG. 6

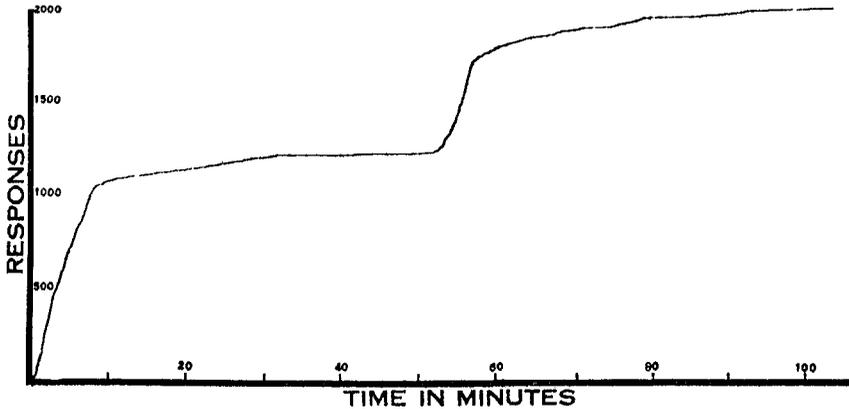


FIG. 7

on a regular schedule—say, every minute—the organism soon forms a discrimination. Little or no responding occurs just after reinforcement, since stimulation from eating is correlated with absence of subsequent reinforcement. How rapidly the discrimination may develop is shown in Fig. 6, which reproduces the first five curves obtained from a pigeon under periodic reinforcement in experimental periods of fifteen minutes each. In the fifth period (or after about one hour of periodic reinforcement) the discrimination yields a pause after each reinforcement, resulting in a markedly stepwise curve. As a result of this discrimination the bird is almost always responding rapidly when reinforced. This is the basis for another discrimination. Rapid responding becomes a favorable stimulating condition. A good example of the effect upon the subsequent extinction curve is shown in Fig. 7. This pigeon had been reinforced once every minute during daily experimental periods of fifteen minutes each for several weeks. In the extinction curve shown, the bird begins to respond at the rate prevailing under the preceding schedule. A quick positive acceleration at the start is lost in the reduction of the record. The

pigeon quickly reaches and sustains a rate that is higher than the overall-rate during periodic reinforcement. During this period the pigeon creates a stimulating condition previously optimally correlated with reinforcement. Eventually, as some sort of exhaustion intervenes, the rate falls off rapidly to a much lower but fairly stable value and then to practically zero. A condition then prevails under which a response is not normally reinforced. The bird is therefore not likely to begin to respond again. When it does respond, however, the situation is slightly improved and, if it continues to respond, the conditions rapidly become similar to those under which reinforcement has been received. Under this "autocatalysis" a high rate is quickly reached, and more than 500 responses are emitted in a second burst. The rate then declines quickly and fairly smoothly, again to nearly zero. This curve is not by any means disorderly. Most of the curvature is smooth. But the burst of responding at forty-five minutes shows a considerable residual strength which, if extinction were merely exhaustion, should have appeared earlier in the curve. The curve may be reasonably accounted for by assuming that the

bird is largely controlled by the preceding spurious correlation between reinforcement and rapid responding.

This assumption may be checked by constructing a schedule of reinforcement in which a differential contingency between rate of responding and reinforcement is impossible. In one such schedule of what may be called "aperiodic reinforcement" one interval between successive reinforced responses is so short that no unreinforced responses intervene while the longest interval is about two minutes. Other intervals are distributed arithmetically between these values, the average remaining one minute. The intervals are roughly randomized to compose a program of reinforcement. Under this program the probability of reinforcement does not change with respect to previous reinforcements, and the curves never acquire the stepwise character of curve E in Fig. 6. (Fig. 9 shows curves from a similar program.) As a result no correlation between different rates of responding and different probabilities of reinforcement can develop.

An extinction curve following a brief exposure to aperiodic reinforcement is shown in Fig. 8. It begins characteristically at the rate prevailing under aperiodic reinforcement and, unlike the curve following regular periodic rein-

forcement, does not accelerate to a higher overall rate. There is no evidence of the "autocatalytic" production of an optimal stimulating condition. Also characteristically, there are no significant discontinuities or sudden changes in rate in either direction. The curve extends over a period of eight hours, as against not quite two hours in Fig. 7, and seems to represent a single orderly process. The total number of responses is higher, perhaps because of the greater time allowed for emission. All of this can be explained by the single fact that we have made it impossible for the pigeon to form a pair of discriminations based, first, upon stimulation from eating and, second, upon stimulation from rapid responding.

Since the longest interval between reinforcement was only two minutes, a certain novelty must still have been introduced as time passed. Whether this explains the curvature in Fig. 8 may be tested to some extent with other programs of reinforcement containing much longer intervals. A geometric progression was constructed by beginning with 10 seconds as the shortest interval and repeatedly multiplying through by a ratio of 1.54. This yielded a set of intervals averaging 5 minutes, the longest of which was more than 21 minutes. Such a set was randomized in a program

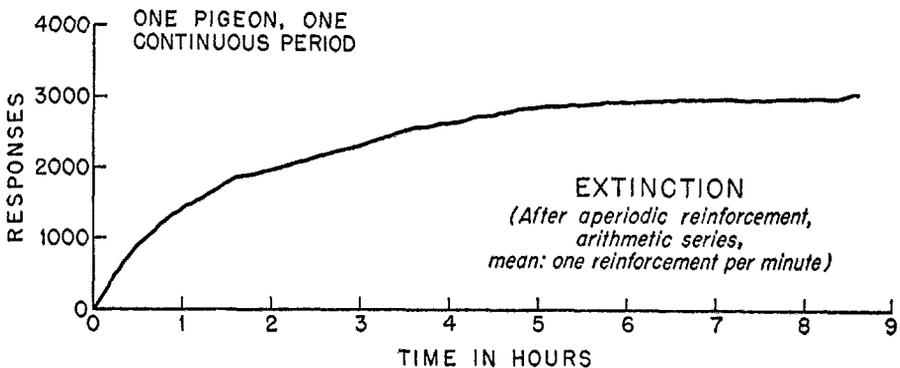


FIG. 8

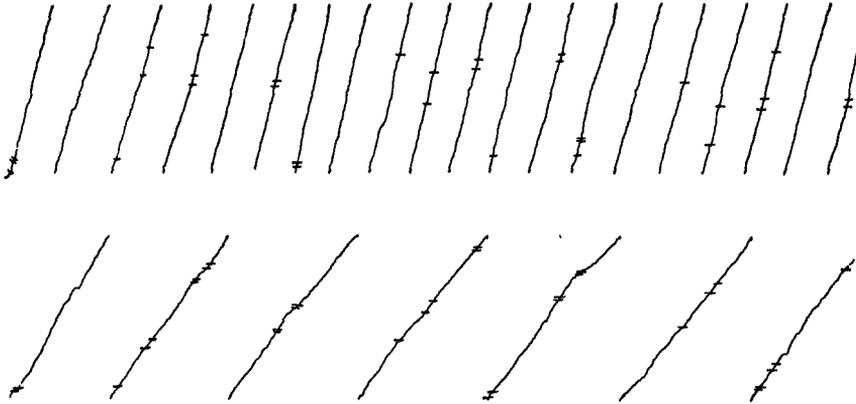


FIG. 9

of reinforcement repeated every hour. In changing to this program from the arithmetic series, the rates first declined during the longer intervals, but the pigeons were soon able to sustain a constant rate of responding under it. Two records in the form in which they were recorded are shown in Fig. 9. (The pen resets to zero after every thousand responses. In order to obtain a single cumulative curve it would be necessary to cut the record and to piece the sections together to yield a continuous line. The raw form may be reproduced with less reduction.) Each reinforcement is represented by a horizontal dash. The time covered is about 3 hours. Records are shown for two pigeons that maintained different overall rates under this program of reinforcement.

Under such a schedule a constant rate of responding is sustained for at least 21 minutes without reinforcement, after which a reinforcement is received. Less novelty should therefore develop during succeeding extinction. In Curve 1 of Fig. 10 the pigeon had been exposed to several sessions of several hours each with this geometric set of intervals. The number of responses emitted in extinction is about twice that of the curve in Fig. 8 after the arithmetic set of in-

tervals averaging one minute, but the curves are otherwise much alike. Further exposure to the geometric schedule builds up longer runs during which the rate does not change significantly. Curve 2 followed Curve 1 after two and one-half hours of further aperiodic reinforcement. On the day shown in Curve 2 a few aperiodic reinforcements were first given, as marked at the beginning of the curve. When reinforcement was discontinued, a fairly constant rate of responding prevailed for several thousand responses. After another experimental session of two and one-half hours with the geometric series, Curve 3 was recorded. This session also began with a short series of aperiodic reinforcements, followed by a sustained run of more than 6000 unreinforced responses with little change in rate (A). There seems to be no reason why other series averaging perhaps more than five minutes per interval and containing much longer exceptional intervals would not carry such a straight line much further.

In this attack upon the problem of extinction we create a schedule of reinforcement which is so much like the conditions that will prevail during extinction that no decline in rate takes

place for a long time. In other words we generate extinction with no curvature. Eventually some kind of exhaustion sets in, but it is not approached gradually. The last part of Curve 3 (unfortunately much reduced in the figure) may possibly suggest exhaustion in the slight overall curvature, but it is a small part of the whole process. The record is composed mainly of runs of a few hundred responses each, most of them at approximately the same rate as that maintained under periodic reinforcement. The pigeon stops abruptly; when it starts to respond again, it quickly reaches the rate of responding under which it was reinforced. This recalls the spurious correlation between rapid responding and reinforcement un-

der regular reinforcement. We have not, of course, entirely eliminated this correlation. Even though there is no longer a differential reinforcement of high against low rates, practically all reinforcements have occurred under a constant rate of responding.

Further study of reinforcing schedules may or may not answer the question of whether the novelty appearing in the extinction situation is entirely responsible for the curvature. It would appear to be necessary to make the conditions prevailing during extinction identical with the conditions prevailing during conditioning. This may be impossible, but in that case the question is academic. The hypothesis, meanwhile, is not a theory in the present sense,

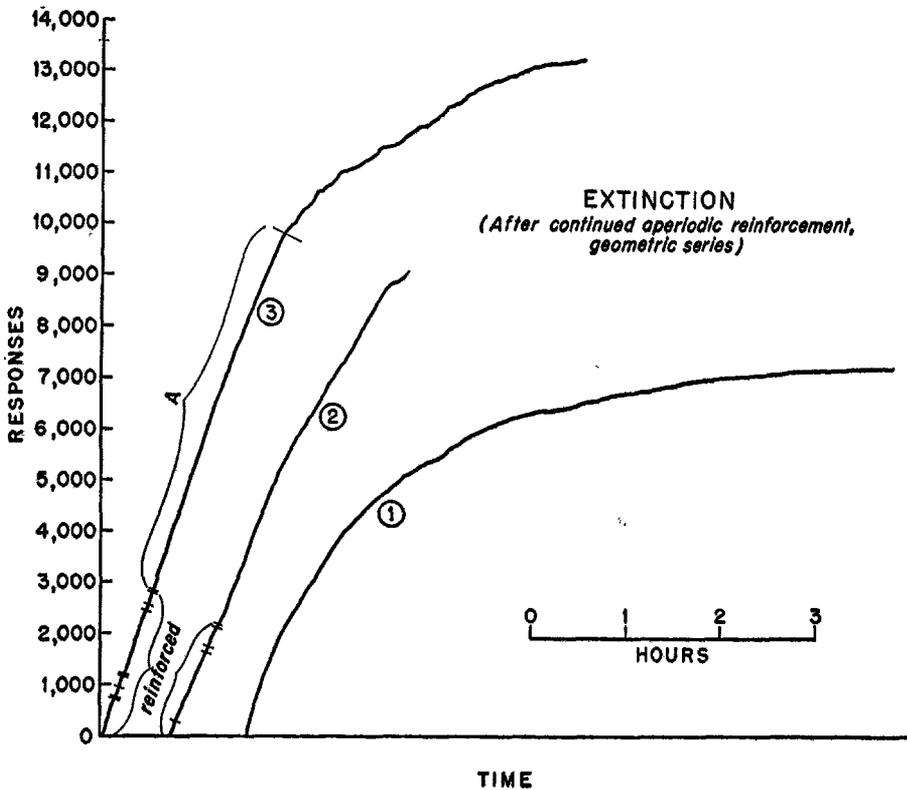


FIG. 10

since it makes no statements about a parallel process in any other universe of discourse.⁴

The study of extinction after different schedules of aperiodic reinforcement is not addressed wholly to this hypothesis. The object is an economical description of the conditions prevailing during reinforcement and extinction and of the relations between them. In using rate of responding as a basic datum we may appeal to conditions that are observable and manipulable and we may express the relations between them in objective terms. To the extent that our datum makes this possible, it reduces the need for theory. When we observe a pigeon emitting 7000 responses at a constant rate without reinforcement, we are not likely to explain an extinction curve containing perhaps a few hundred responses by appeal to the piling up of reaction inhibition or any other fatigue product. Research which is conducted without commitment to theory is more likely to carry the study of extinction into new areas and new orders of magnitude. By hastening the accumulation of data, we speed the departure of theories. If the theories have played no part in the design of our experiments, we need not be sorry to see them go.

Complex Learning

A third type of learning theory is illustrated by terms like *preferring*, *choosing*, *discriminating*, and *matching*. An effort may be made to define these solely in terms of behavior, but in traditional practice they refer to processes in another dimensional system. A re-

⁴ It is true that it appeals to stimulation generated in part by the pigeon's own behavior. This may be difficult to specify or manipulate, but it is not theoretical in the present sense. So long as we are willing to assume a one-to-one correspondence between action and stimulation, a physical specification is possible.

sponse to one of two available stimuli may be called choice, but it is commoner to say that it is the result of choice, meaning by the latter a theoretical pre-behavioral activity. The higher mental processes are the best examples of theories of this sort; neurological parallels have not been well worked out. The appeal to theory is encouraged by the fact that choosing (like discriminating, matching, and so on) is not a particular piece of behavior. It is not a response or an act with specified topography. The term characterizes a larger segment of behavior in relation to other variables or events. Can we formulate and study the behavior to which these terms would usually be applied without recourse to the theories which generally accompany them?

Discrimination is a relatively simple case. Suppose we find that the probability of emission of a given response is not significantly affected by changing from one of two stimuli to the other. We then make reinforcement of the response contingent upon the presence of one of them. The well-established result is that the probability of response remains high under this stimulus and reaches a very low point under the other. We say that the organism now discriminates between the stimuli. But discrimination is not itself an action, or necessarily even a unique process. Problems in the field of discrimination may be stated in other terms. How much induction obtains between stimuli of different magnitudes or classes? What are the smallest differences in stimuli that yield a difference in control? And so on. Questions of this sort do not presuppose theoretical activities in other dimensional systems.

A somewhat larger segment must be specified in dealing with the behavior of choosing one of two concurrent stimuli. This has been studied in the pigeon by examining responses to two keys differ-

ing in position (right or left) or in some property like color randomized with respect to position. By occasionally reinforcing a response on one key or the other without favoring either key, we obtain equal rates of responding on the two keys. The behavior approaches a simple alternation from one key to the other. This follows the rule that tendencies to respond eventually correspond to the probabilities of reinforcement. Given a system in which one key or the other is occasionally connected with the magazine by an external clock, then if the right key has just been struck, the probability of reinforcement *via* the left key is higher than that *via* the right since a greater interval of time has elapsed during which the clock may have closed the circuit to the left key. But the bird's behavior does not correspond to this probability merely out of respect for mathematics. The specific result of such a contingency of reinforcement is that changing-to-the-other-key-and-striking is more often reinforced than striking-the-same-key-a-second-time. We are no longer dealing with just two responses. In order to analyze "choice" we must consider a single final response, striking, without respect to the position or color of the key, and in addition the responses of changing from one key or color to the other.

Quantitative results are compatible with this analysis. If we periodically reinforce responses to the right key only, the rate of responding on the right will rise while that on the left will fall. The response of changing-from-right-to-left is never reinforced while the response of changing-from-left-to-right is occasionally so. When the bird is striking on the right, there is no great tendency to change keys; when it is striking on the left, there is a strong tendency to change. Many more responses come to be made to the right

key. The need for considering the behavior of changing over is clearly shown if we now reverse these conditions and reinforce responses to the left key only. The ultimate result is a high rate of responding on the left key and a low rate on the right. By reversing the conditions again the high rate can be shifted back to the right key. In Fig. 11 a group of eight curves have been averaged to follow this change during six experimental periods of 45 minutes each. Beginning on the second day in the graph responses to the right key (R^R) decline in extinction while responses to the left key (R^L) increase through periodic reinforcement. The mean rate shows no significant varia-

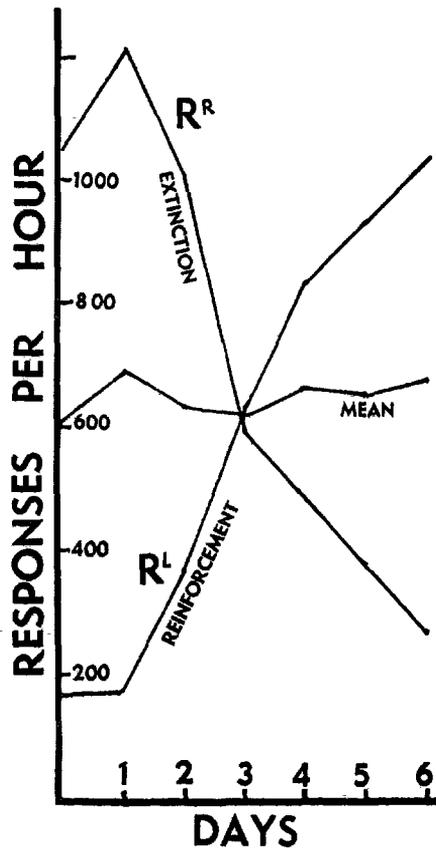


FIG. 11

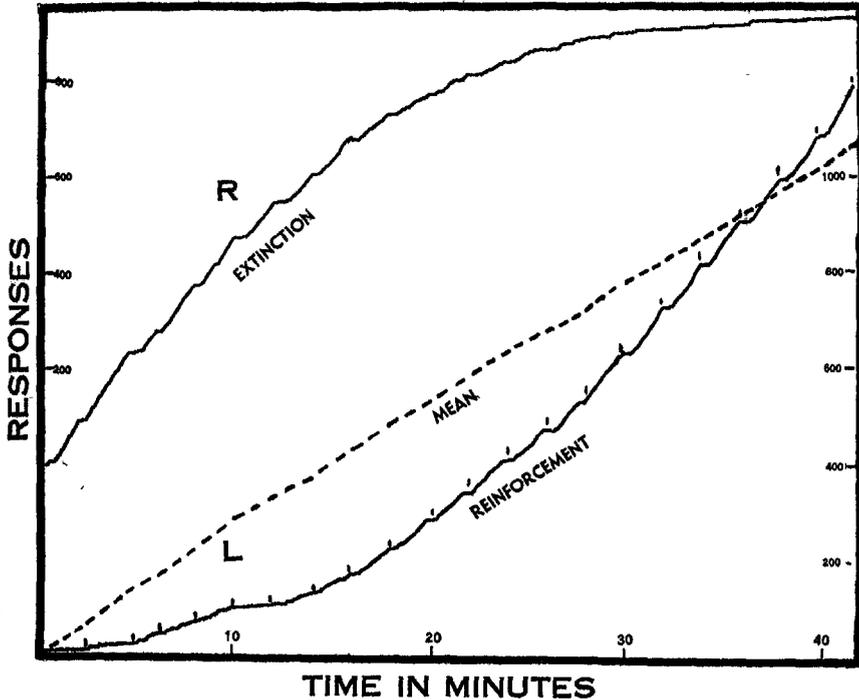


FIG. 12

tion, since periodic reinforcement is continued on the same schedule. The mean rate shows the condition of strength of the response of striking a key regardless of position. The distribution of responses between right and left depends upon the relative strength of the responses of changing over. If this were simply a case of the extinction of one response and the concurrent reconditioning of another, the mean curve would not remain approximately horizontal since reconditioning occurs much more rapidly than extinction.⁵

The rate with which the bird changes from one key to the other depends upon the distance between the keys. This distance is a rough measure of the stimulus-difference between the two keys. It

⁵ Two topographically independent responses, capable of emission at the same time and hence not requiring change-over, show separate processes of reconditioning and extinction, and the combined rate of responding varies.

also determines the scope of the response of changing-over, with an implied difference in sensory feedback. It also modifies the spread of reinforcement to responses supposedly not reinforced, since if the keys are close together, a response reinforced on one side may occur sooner after a preceding response on the other side. In Fig. 11 the two keys were about one inch apart. They were therefore fairly similar with respect to position in the experimental box. Changing from one to the other involved a minimum of sensory feedback, and reinforcement of a response to one key could follow very shortly upon a response to the other. When the keys are separated by as much as four inches, the change in strength is much more rapid. Fig. 12 shows two curves recorded simultaneously from a single pigeon during one experimental period of about 40 minutes. A high rate

to the right key and a low rate to the left had previously been established. In the figure no responses to the right were reinforced, but those to the left were reinforced every minute as indicated by the vertical dashes above curve L. The slope of R declines in a fairly smooth fashion while that of L increases, also fairly smoothly, to a value comparable to the initial value of R. The bird has conformed to the changed contingency within a single experimental period. The mean rate of responding is shown by a dotted line, which again shows no significant curvature.

What is called "preference" enters into this formulation. At any stage of the process shown in Fig. 12 preference might be expressed in terms of the relative rates of responding to the two keys. This preference, however, is not in striking a key but in changing from one key to the other. The probability that the bird will strike a key regardless of its identifying properties behaves independently of the preferential response of changing from one key to the other. Several experiments have revealed an additional fact. A preference remains fixed if reinforcement is withheld. Fig. 13 is an example. It shows simultaneous extinction curves from two keys during seven daily experimental periods of one hour each. Prior to extinction the relative strength of the responses of changing-to-R and changing-to-L yielded a "preference" of about 3 to 1

for R. The constancy of the rate throughout the process of extinction has been shown in the figure by multiplying L through by a suitable constant and entering the points as small circles on R. If extinction altered the preference, the two curves would not be superimposed in this way.

These formulations of discrimination and choosing enable us to deal with what is generally regarded as a much more complex process—matching to sample. Suppose we arrange three translucent keys, each of which may be illuminated with red or green light. The middle key functions as the sample and we color it either red or green in random order. We color the two side keys one red and one green, also in random order. The "problem" is to strike the side key which corresponds in color to the middle key. There are only four three-key patterns in such a case, and it is possible that a pigeon could learn to make an appropriate response to each pattern. This does not happen, at least within the temporal span of the experiments to date. If we simply present a series of settings of the three colors and reinforce successful responses, the pigeon will strike the side keys without respect to color or pattern and be reinforced 50 per cent of the time. This is, in effect, a schedule of "fixed ratio" reinforcement which is adequate to maintain a high rate of responding.

Nevertheless it is possible to get a pigeon to match to sample by reinforcing the discriminative responses of striking-red-after-being-stimulated-by-red and striking-green-after-being-stimulated-by-green while extinguishing the other two possibilities. The difficulty is in arranging the proper stimulation at the time of the response. The sample might be made conspicuous—for example, by having the sample color in the general illumination of the experimental box. In such a case the pi-

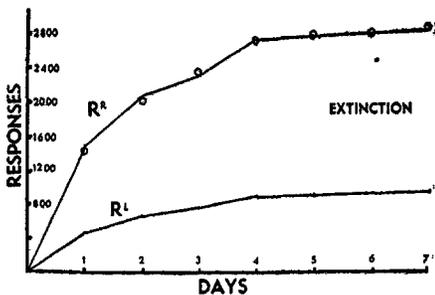


FIG. 13

geon would learn to strike red keys in a red light and green keys in a green light (assuming a neutral illumination of the background of the keys). But a procedure which holds more closely to the notion of matching is to induce the pigeon to "look at the sample" by means of a separate reinforcement. We may do this by presenting the color on the middle key first, leaving the side keys uncolored. A response to the middle key is then reinforced (secondarily) by illuminating the side keys. The pigeon learns to make two responses in quick succession—to the middle key and then to one side key. The response to the side key follows quickly upon the visual stimulation from the middle key, which is the requisite condition for a discrimination. Successful matching was readily established in all ten pigeons tested with this technique. Choosing the opposite is also easily set up. The discriminative response of striking-red-after-being-stimulated-by-red is apparently no easier to establish than striking-red-after-being-stimulated-by-green. When the response is to a key of the same color, however, generalization may make it possible for the bird to match a new color. This is an extension of the notion of matching that has not yet been studied with this method.

Even when matching behavior has been well established, the bird will not respond correctly if all three keys are now presented at the same time. The bird does not possess strong behavior of looking at the sample. The experimenter must maintain a separate reinforcement to keep this behavior in strength. In monkeys, apes, and human subjects the ultimate success in choosing is apparently sufficient to reinforce and maintain the behavior of looking at the sample. It is possible that this species difference is simply a difference in the temporal relations required for reinforcement.

The behavior of matching survives unchanged when all reinforcement is withheld. An intermediate case has been established in which the correct matching response is only periodically reinforced. In one experiment one color appeared on the middle key for one minute; it was then changed or not changed, at random, to the other color. A response to this key illuminated the side keys, one red and one green, in random order. A response to a side key cut off the illumination to both side keys, until the middle key had again been struck. The apparatus recorded all matching responses on one graph and all non-matching on another. Pigeons which have acquired matching behavior under continuous reinforcement have maintained this behavior when reinforced no oftener than once per minute on the average. They may make thousands of matching responses per hour while being reinforced for no more than sixty of them. This schedule will not necessarily develop matching behavior in a naive bird, for the problem can be solved in three ways. The bird will receive practically as many reinforcements if it responds to (1) only one key or (2) only one color, since the programming of the experiment makes any persistent response eventually the correct one.

A sample of the data obtained in a complex experiment of this sort is given in Fig. 14. Although this pigeon had learned to match color under continuous reinforcement, it changed to the spurious solution of a color preference under periodic reinforcement. Whenever the sample was red, it struck both the sample and the red side key and received all reinforcements. When the sample was green, it did not respond and the side keys were not illuminated. The result shown at the beginning of the graph in Fig. 14 is a high rate of responding on the upper graph, which records match-

ing responses. (The record is actually step-wise, following the presence or absence of the red sample, but this is lost in the reduction in the figure.) A color preference, however, is not a solution to the problem of opposites. By changing to this problem, it was possible to change the bird's behavior as shown between the two vertical lines in the figure. The upper curve between these lines shows the decline in matching responses which had resulted from the color preference. The lower curve between the same lines shows the development of responding to and matching the opposite color. At the second vertical line the reinforcement was again made contingent upon matching. The upper curve shows the reestablishment of matching behavior while the lower curve shows a decline in striking the opposite color. The result was a true solution: the pigeon struck the sample, no matter what its color, and then the corresponding side key. The lighter line connects the means of a series of points

on the two curves. It seems to follow the same rule as in the case of choosing: changes in the distribution of responses between two keys do not involve the over-all rate of responding to a key. This mean rate will not remain constant under the spurious solution achieved with a color preference, as at the beginning of this figure.

These experiments on a few higher processes have necessarily been very briefly described. They are not offered as proving that theories of learning are not necessary, but they may suggest an alternative program in this difficult area. The data in the field of the higher mental processes transcend single responses or single stimulus-response relationships. But they appear to be susceptible to formulation in terms of the differentiation of concurrent responses, the discrimination of stimuli, the establishment of various sequences of responses, and so on. There seems to be no *a priori* reason why a complete account is not possible without appeal to theoretical processes in other dimensional systems.

Conclusion

Perhaps to do without theories altogether is a *tour de force* that is too much to expect as a general practice. Theories are fun. But it is possible that the most rapid progress toward an understanding of learning may be made by research that is not designed to test theories. An adequate impetus is supplied by the inclination to obtain data showing orderly changes characteristic of the learning process. An acceptable scientific program is to collect data of this sort and to relate them to manipulable variables, selected for study through a common sense exploration of the field.

This does not exclude the possibility of theory in another sense. Beyond the collection of uniform relationships lies

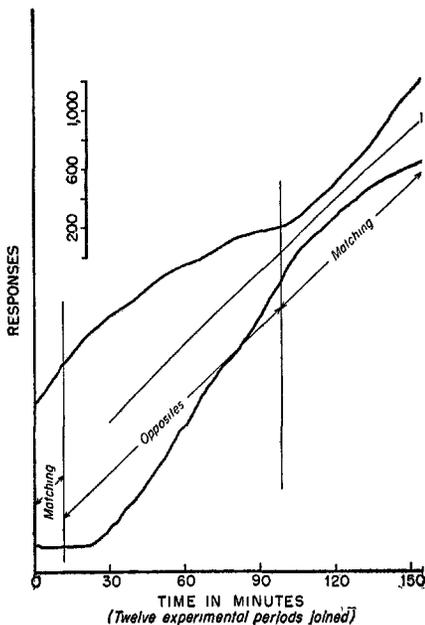


FIG. 14

the need for a formal representation of the data reduced to a minimal number of terms. A theoretical construction may yield greater generality than any assemblage of facts. But such a construction will not refer to another dimensional system and will not, therefore, fall within our present definition. It will not stand in the way of our search for functional relations because it will arise only after relevant variables have been found and studied. Though it may be difficult to understand, it will not be easily misunderstood, and it will have none of the objectionable effects of the theories here considered.

We do not seem to be ready for theory in this sense. At the moment we make little effective use of empirical, let alone rational, equations. A few of the present curves could have been fairly closely fitted. But the most elementary preliminary research shows that

there are many relevant variables, and until their importance has been experimentally determined, an equation that allows for them will have so many arbitrary constants that a good fit will be a matter of course and a cause for very little satisfaction.

REFERENCES

1. MOWRER, O. H., & JONES, H. M. Extinction and behavior variability as functions of effortfulness of task. *J. exp. Psychol.*, 1943, 33, 369-386.
2. SKINNER, B. F. *The behavior of organisms*. New York: D. Appleton-Century Co., 1938.
3. ——. The nature of the operant reserve. *Psychol. Bull.*, 1940, 37, 423 (abstract).
4. ——. Differential reinforcement with respect to time. *Amer. Psychol.*, 1946, 1, 274-275 (abstract).
5. ——. The effect of the difficulty of a response upon its rate of emission. *Amer. Psychol.*, 1946, 1, 462 (abstract).

[MS. received December 5, 1949]