UNIVERSAL LIBRARY

Learning Theory, Personality Theory, and Clinical Research

Eleven lectures
given under the auspices of
the Department of Psychology in
the College of Arts and Sciences
of the University of Kentucky
on March 13 and 14, 1953

Learning Theory, Personality Theory, and Clinical Research

THE KENTUCKY SYMPOSIUM

DONALD K. ADAMS

R. B. AMMONS

JOHN M. BUTLER

RAYMOND B. CATTELL

HARRY F. HARLOW

NORMAN R. F. MAIER

O. H. MOWRER

DONALD SNYGG

KENNETH W. SPENCE

DELOS D. WICKENS

J. R. WITTENBORN

NEW YORK · JOHN WILEY & SONS, INC.

LONDON · CHAPMAN & HALL, LIMITED

PRINTED IN THE UNITED STATES OF AMERICA

Foreword

In recent years learning theory, personality theory, and clinical research have been among the most vigorously cultivated fields in American psychology. Many workers in these fields have tried to integrate the results of their labor. Some learning theorists have devoted themselves to problems which seem important also for personality theory and clinical psychology. Since the aim of psychotherapy is to produce changes, and the study of such changes is the study of learning, some clinicians interested in psychotherapy have looked to learning theory for help in explaining the development of normal and abnormal personalities and for guidance in building up a theory of psychotherapy.

In order to encourage current tendencies leading toward a closer integration of these three branches of psychology, the Department of Psychology of the University of Kentucky decided in the autumn of 1952 to hold a symposium on the relationships among these three areas. The original proposal was made by Dr. Robert E. Bills, who was made Chairman of a Symposium Committee, the other members of which were Drs. Lysle W. Croft, P. L. Mellenbruch, Robert D. North, and Harold Webster. The University administration gave its cordial support to the project, and a number of distinguished psychologists accepted invitations to contribute lectures.

The papers here published were presented at the Symposium on March 13 and 14, 1953, before members of the University community and a considerable number of visitors.

We wish to express our thanks to the President of the University, Dr. Herman L. Donovan, for his interest and for his address of welcome at the opening meeting. We also wish to thank Dr. James S. Calvin, Dr. Betsy W. Estes, Dr. Charles F. Diehl, Dr. Ernest Meyers, Professor Edward Newbury, and Mrs. Lysle W. Croft for their contributions, made in a variety of ways, to the success of the Symposium. Dr. Frank A. Pattie was asked to assume the task of publication of the Symposium, and we wish to express our appreciation of his work in finding a publisher and in seeing the book through the press.

THE SYMPOSIUM COMMITTEE

Lexington, Kentucky,
February 12, 1954

Contributors

Donald K. Adams	Professor of Psychology, Duke University
R. B. Ammons	Assistant Professor of Psychology, University of Louisville
John M. Butler	Assistant Professor of Psychology, University of Chicago
Raymond B. Cattell	Research Professor of Psychology, University of Illinois
Harry F. Harlow	Professor of Psychology, University of Wisconsin
Norman R. F. Maier	Professor of Psychology, University of Michigan
O. H. Mowrer	Research Professor of Psychology, University of Illinois
Donald Snygg	Professor and Chairman of the Department of Psychology, State University of New York Teachers College, Oswego
Kenneth W. Spence	Professor and Head of the Department of Psychology, State University of Iowa
Delos D. Wickens	Professor of Psychology, Ohio State University
J. R. Wittenborn	Research Associate and Associate Professor, Department of Psychology, Yale University

Contents

Current Interpretations of Learning Data and Some Recent Developments in Stimulus-Response Theory

KENNETH W. SPENCE

In this opening address of the Symposium on Relationships between Learning Theory, Personality Theory, and Clinical Research my assignment is twofold: first, to discuss briefly the data investigated in learning experiments and the theories proposed concerning them, and second, to present some recent developments in my own theoretical position. The limited time available necessitates a highly selective treatment of both topics. Hence I have rather arbitrarily chosen to emphasize certain aspects of them on the basis either of their interest to me at the moment or of their seeming relevance to this symposium.

I

The first point that I should like to bring up has to do with a matter that is of considerable importance in any discussion of learning theory, particularly one that occurs in the context of such topics as personality theory and clinical psychology. I have reference to the importance, indeed the necessity, for us to have clearly before us what it is that learning theories are about. What are the phenomena that the learning psychologist has taken as the object of his studies and about which he has attempted to formulate his theories?

All of us, of course, are more or less familiar in a general way with what is meant by learning phenomena, and we have each actually had many first-hand experiences with instances of learning in everyday life. It should be noted, however, that these familiar instances do not constitute the data with which the learning psychologist has been concerned. The body of laws and theories that he has developed has its origins in a very different set of observations. These observations have involved a variety of laboratory arrangements in which such things as the lever-pressing behavior and running speed of the white rat, the salivary response of dogs, and the closure of the human eyelid have constituted the observed phenomena. The laws that the experimental learning psychologist has discovered and the theories that he has formulated with respect to them grew out of these laboratory phenomena and as yet have not been related to instances of learning in "real" life.

Unfortunately, the fact that learning psychologists deliberately chose not to concern themselves with "real" life situations has not always been realized. Still more unfortunate is the fact that, when there has been appreciation of it, considerable misunderstanding of the purpose of working with such phenomena has existed. Why, the question is frequently asked, do not learning psychologists concentrate their efforts on life situations and use human subjects instead of white rats, dogs, and monkeys?

There are a number of answers to such complaints, the complete consideration of which, however, would take us too far astray from our main concern. My comments will be confined to just two points. (1) It should be noted that there are psychologists who do concern themselves with learning behavior as it occurs in everyday life. Such phenomena have been and are being studied by a number of different psychologists, e.g., by the child psychologist, the educational and social psychologists, and the clinical psychologist. Presumably whatever knowledge can be gained by such more or less naturalistic, uncontrolled observation is being obtained. That such studies have not led us very far in the attainment of precise laws of even the lowest-order type is pretty well known to everyone, and it is this fact that brings me to my second point. (2) From the many examples in the history of the development of knowledge in other fields of science we have come, or at least we should have come, to appreciate that the first objective of the scientist, assuming that he has developed adequate techniques of observation, is to discover the low-order, empirical laws holding among the particular variables specified by his measuring techniques. Both the physical and biological scientists have demonstrated over and over the importance of experimental control and systematic variation of the combinations of factors in the situation under observation if these laws are to be discovered. Furthermore, these scientists did not hesitate, as some psychologists have, to arrange unreal (i.e., unworld-like) situations in order to achieve this goal. Although nature may abhor a vacuum, fortunately the experimental physicist did not. Likewise, the biologist never troubled himself about the fact that the isolated piece of tissue in his test tube was unlike anything in real life. It may be noted, moreover, that the laws such artificial arrangements helped to discover and formulate subsequently were found to provide for the explanation of events that do occur under natural conditions.

It is in this tradition that the experimental learning psychologist has worked. Interested in providing the same kind of knowledge about behavior phenomena that scientists in other fields have developed, he has proceeded in a comparable manner to arrange for the successive isolation of various aspects of learning phenomena so that their interrelations might be ascertained. Thus in the case of the individual organism he has attempted in some of his experiments to eliminate the operation of certain not-too-well-known and hence uncontrollable processes that we refer to as thinking or reasoning. One manner of doing this has been to use non-articulate organisms such as the white rat. Another has been to employ situations such as the memory drum set-up in which the temporal sequence of stimuli and responses is such as to preclude the operation of these unobservable processes. In a similar manner the learning psychologist has attempted to isolate and control the various aspects of the environment eliciting and following the responses of the organism. Finally, control and manipulation of the motivational factors have also been an important part of the learning psychologist's experimental efforts.

As the result of these controlled laboratory studies a number of important behavioral and environmental variables have been identified, and at least a good beginning has been made in discovering and formulating the type of lawful relations among them that scientists seek. For the most part, it is true, these laws are of the low-order type involving specific response variables in the several specific experimental situations (classical conditioning, discrimination learning, paired associate learning, etc.). As yet only a start has been made toward the development of the type of higher-order generalizations (i.e., theories) that serve to integrate or unify the laws specific to each experimental situation.

At this point we need to take a little closer look at the types of laws, hypotheses, and theories that the experimental studies of the learning psychologist have produced so far. The laws provided by the different kinds of learning experiments consist in a set of empirical functions which relate the various response measures to a number of manipulable environmental variables. The following represent some of these investigated relationships for one response measure (e.g., frequency of response) in one experimental situation (classical conditioning):

1. $R = f$(Number of trials, N).
2. $R = f$(Intensity of the conditioned stimulus, S_C).
3. $R = f$(Intensity of the unconditioned stimulus, S_U).
4. $R = f$(Time interval between S_C and S_U — $T_{S_c-S_u}$).
5. $R = f$(Time between successive trials, T_R).
6. $R = f$(Amount of work involved in R, W).

   $R = f(N, S_C, S_U, T_{S_c-S_u}, T_R, W, \text{etc.})$.

The so-called learning function (No. 1 in the list) is one of these laws. Although it is the function in which learning psychologists have been most interested, the relations of performance measures to such other variables as time of food deprivation (T_D), delay of reward (T_G), and intensity of the UCS (S_U) are also important for a complete account of the behavior of the subject in the situation.

A number of points concerning these empirical laws are of interest to us. First is the experimental fact that a law involving a particular response measure may be very different from one experimental situation to another. For example, the percentage of correct responses as a function of successive blocks of training trials is negatively accelerated in most simple T-maze studies, whereas the same measure shows a period of initial acceleration in difficult discrimination situations. Likewise, the learning functions for different response measures in the same situation may take very different forms. Thus the frequency measure typically shows an S-shaped curve in classical conditioning, whereas resistance to extinction is usually a negatively accelerated function. In other words, the laws found in our studies are to a very considerable extent highly specific to each experimental situation and to each response measure. We shall return later to a discussion of the implication of this fact for learning theory.

A second point in connection with these empirical laws of behavior in learning situations is that, in addition to being interested in the relation holding between each response measure and each independently manipulable environmental variable, the psychologist is also very much concerned with the problem of how these latter variables combine or interact with each other to determine response strength. In other words, he is interested in determining the precise nature of the final complex function for each response measure in each experimental situation.

With these two aspects of the laws of behavior in learning situations before us, let us now turn to a brief review of the kinds of theories that have come out of these learning experiments. As we shall see, there are a number of very different things that are called theories, and theories serve a number of different purposes.

One of the more prominent kinds of learning theory has been the intervening-variable type introduced by Tolman [*26*, *27*]. Conceiving one of the main tasks of the learning psychologist at the present time to be that of discovering and formulating the precise nature of such complex laws as that shown at the bottom of the list on p. 4, Tolman proposed that it was necessary to introduce theoretical constructs or, as he termed them, intervening variables, in order to do so. He believed that the complicated nature of the function relating a response variable to its several determining variables made it necessary to proceed by conceiving of the function as being analyzed into successive sets of simpler component functions. According to Tolman, these component functions begin by defining a set of intervening variables in terms of the independent variables. Further intervening variables are then introduced by stating them as functions of the first set, until finally the dependent behavior variable is postulated as some function of one or more of the intervening variables.

Actually Tolman never got around to demonstrating how such theoretical constructs provided us with these complex laws, and more recently he himself seems to have questioned their usefulness [*28*]. While it may or may not be a fact that the discovery of the precise nature of such complex functions is facilitated by this intervening-variable type of theorizing, attention should be called to the fact that it is entirely possible to ascertain them by purely experimental means. Thus one could proceed to study the response variable as a joint function of various combinations of the independent variables by means of the factorial type of experiment. An instance of this kind of experiment is the well-known Perin-Williams study [*14*], which showed that the variables, time of deprivation and number of trials, combine in a multiplicative function to determine response strength. We have recently obtained data in the Iowa laboratory which show that the UCS in classical conditioning and number of trials likewise combine in a multiplicative manner to determine response strength. The main point here, however, is that it is not *necessary* to introduce intervening variables to be able to formulate such complex laws. The problem is, rather, an experimental one, one that we ought to be working at more than we are.

A second type of theory that has also employed the intervening-variable type of construct is represented by the work of Hull [*9, 10, 11*]. The primary concern of this type of theory, as I have pointed out in a number of previous discussions [*17, 18*], has been to build a theoretical structure or model that would serve to derive and hence interrelate the specific laws found in the different learning situations. In his book *Principles of behavior* Hull began with the empirical findings from two of the simplest learning situations—classical and instrumental conditioning. On the basis of these data he proposed a hypothetical system of laws involving intervening variables and experimental variables that would provide for the derivation of the different laws for each response measure in these two situations. In its initial formulation such a theoretical structure is, it is true, purely *ad hoc*. The theorist makes those particular assumptions that will lead to the derivation of the known empirical laws. Once formulated on the basis of this initial set of data, however, the theoretical model may be tested in terms of the degree to which it is able to derive lawful relations appearing in new data, whether from the same situation or from other learning situations, e.g., in discrimination learning, maze learning, and rote serial learning. Hull's posthumously published book, *A behavior system*, represents an attempt to apply his theoretical model to new, more complex learning situations. For a number of years the writer and his students have been engaged in a similar application of it to the phenomena of discrimination and simple selective learning [*5, 19, 24*]. Other theories that are essentially similar, in principle, to this type are those of Thurstone [*25*], Gulliksen [*6*], Pitts [*15*], and the more recent formulations of Estes [*3*] and Bush and Mosteller [*1*].

A third class of learning theories is represented by the mathematical formulations of Rashevsky and his students [*8, 16*]. These theories, like those of the previous group, also attempt to integrate the laws from the different learning situations into a more comprehensive system of knowledge. They differ, however, in the origin of their mathematical assumptions. Instead of beginning with the data from one or another type of learning situation they start with certain knowledge from neurophysiology. By means of further assumptions as to the neurological circuits involved in the different learning situations they then derive rational equations representative of empirical relationships to be expected in learning data. This type of theory is perhaps of even less interest to the psychologist concerned with the application of learning theory to phenomena of everyday life than the other theories we have been discussing. Its main contribution is the integration it promises to provide of the laws of neurophysiology with the behavioral laws found in the learning psychologist's laboratory.

In a fourth and final class of learning theories I have lumped a variety of widely different kinds of theorizing that do not fall into the first three groups. For the most part these are limited, specific hypotheses rather than the more comprehensive type of theory that attempts to integrate different areas of laws. These hypotheses vary from being mere guesses as to the role some previously unsuspected variable plays in a particular experimental set-up, through speculations as to the physiological mechanism underlying some bit of behavior, to extrapolations of specific laws discovered in controlled experimental situations to analogous situations in everyday life. Miller and Dollard in their books *Social learning and imitation* and *Personality and psychotherapy* provide many examples of this sort of hypothesizing and in so doing have made a start, at least, in tying together the phenomena of everyday learning and the more precise laws discovered in the laboratory. Instances of such limited hypotheses in experimental learning data are the drive-reduction hypothesis of reinforcement and the hypothetical role of the fractional anticipatory goal response in mediating the appropriate response in latent learning experiments.

This concludes the first part of my assignment. In summary, I have emphasized the point that the data heretofore investigated by the theory-oriented learning psychologist are, for the most part, quite different from the observations of learning in everyday life. They consist, as we have seen, in observations obtained under controlled laboratory conditions that often depart radically from those of everyday life. The main characteristics of the laws and the different types of theories that these experimental studies have led to were outlined briefly, and the need for such an approach was indicated if the abstract kind of knowledge that the scientist seeks is to be found.

One final comment before leaving this discussion is concerned with the question as to whether the theories developed from these laboratory experiments, when sufficiently elaborated, will ever provide for the explanation and control of behavior in real life situations. No definitive answer, of course, can be given at the present time, for as yet none of them is sufficiently abstract or complete to account even for all the laboratory findings. One can only point to the success that such types of theory have had in the physical sciences and assume that the same developments will occur in psychology. It is perhaps worth while for a few psychologists, at least, to have this faith, if for no other reason than to provide a control experiment to test the beliefs of the many who know on a priori grounds that psychology is different in this respect from other sciences.

II

In turning to the consideration of some recent developments in my own theoretical notions concerning learning I should, perhaps, begin by identifying my position in terms of the different classes of theories mentioned earlier. As was indicated in my discussion of Hull's theorizing, I have for a number of years been interested in extending the quantitative aspects of this type of theory to more complex learning phenomena, particularly discrimination learning. Through the medium of doctoral theses we have also been engaged in checking and modifying, where necessary, the basic theoretical structure as developed within the context of classical and instrumental conditioning.

This position, as you know, is typically referred to as the *S–R* reinforcement theory, and as such it is usually contrasted with such *S–S* contiguity theories as that of Tolman and even such *S–R* contiguity positions as that represented by Guthrie [7]. Although the issues of *S–S* versus *S–R* and reinforcement versus contiguity are genuine enough ones, they do not in my opinion warrant the amount of attention that has been accorded them. Furthermore, from the point of view of the development of a quantitative theory that will serve to integrate the laws found in different learning situations they are of no immediate importance. In other words, such quantitative theorizing is, or can be, entirely independent of these issues.

Although the concept of *S–R* tendency will be used in the ensuing presentation, no bias is intended as to whether the hypothetical learning change is a connection between two sensory elements of the nervous system or between afferent and efferent units. The essential meaning is that of a quantitative value that changes in some specified amount as the result of certain experimental operations. Likewise the term "reinforcement" will be used but only in the general sense that the stimulus effects or consequences of a response in some way or other are related to the changing value of the *S–R* tendency.

We shall have time to discuss only two recent developments in this quantitative *S–R* theory. The first that we shall consider is not, strictly speaking, a new development, but represents rather an emendation of one portion of Hull's basic theoretical formulation that seems to have escaped the attention of even his most ardent critics. Figure 1 presents the variables, experimental and hypothetical, that are involved in the theory developed for classical and instrumental conditioning. At the left are represented the experimental variables that have been found to affect response strength (S_C, intensity of the CS; $T_{S_C-S_U}$, time interval between onsets of CS and UCS; T_G, time of delay of reward; N, number of reinforced trials, etc.). In the middle are

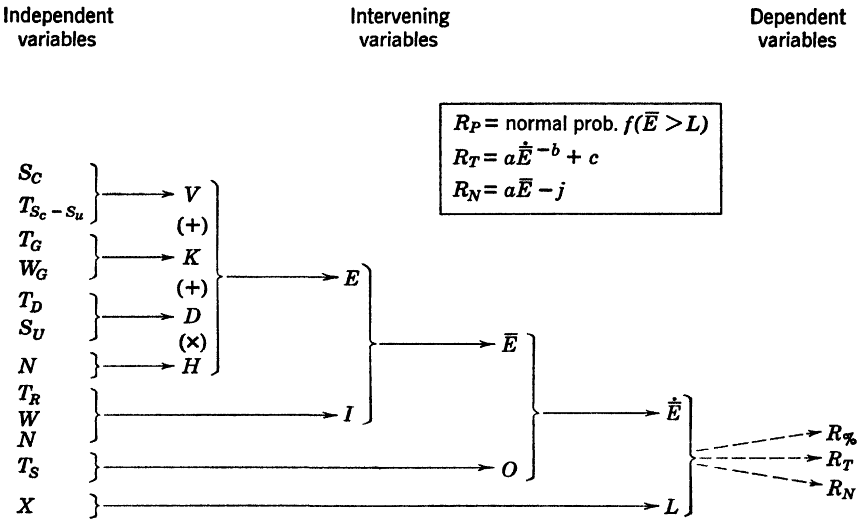

Fig. 1. Diagram showing relation between the experimental variables and the intervening variables in Hull's theory of simple learning phenomena.

the intervening theoretical variables, and at the right are the dependent response variables. The part of this theoretical structure that I wish to discuss now is the last set of relations (represented by the broken arrows) at the extreme right of the figure.

In the *Principles of behavior* Hull called attention to the importance of anchoring the intervening theoretical constructs both to the antecedent environmental conditions and to the consequent response events. The latter he proposed to accomplish by means of postulates relating the different response measures to one or another of the relevant theoretical constructs. Thus, in the case of the three response measures we shall be considering, he made the specific assumptions shown at the top right of Fig. 1.

1. The probability of occurrence of a response (R_P) is a normal, integral function of the amount that the effective excitatory potential ($\bar{E}$) exceeds the reaction threshold (L).

2. The latency of response (R_T) is a decreasing hyperbolic function of the momentary effective excitatory potential ($\dot{\bar{E}}$).

3. The number of trials required to produce extinction (R_N) is a simple linear increasing function of the effective excitatory potential ($\bar{E}$).

Actually these postulated relations are entirely unnecessary, for it is possible to derive relations between these response measures and one or another of the intervening theoretical constructs on the basis of already existing postulates. Thus it may be shown that the func-

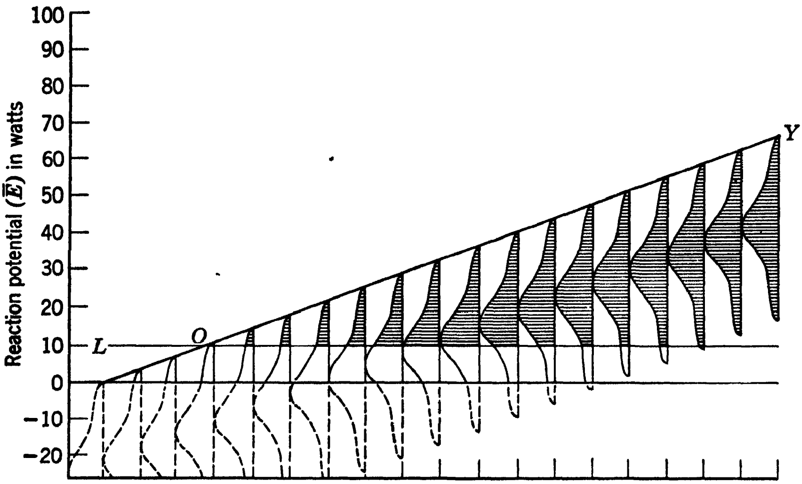

Fig. 2. Diagram showing relations between intervening variables, effective excitatory potential ($\bar{E}$), oscillatory inhibition (O), and reaction threshold (L).

tion relating probability of response (R_P) to $\bar{E}$ may be derived from the assumptions already made concerning oscillatory inhibitory potential (O) and the response threshold (L). This derivation may be shown graphically by means of Fig. 2. In this graph a linear increase in $\bar{E}$ is represented by the line OY; L represents the threshold value that a particular momentary value of $\bar{E}$ must exceed for a response to occur. The upended normal distributions represent the frequency of occurrence of the oscillatory O values that are subtracted from effective excitatory potential ($\bar{E}$) to give the momentary effective excitatory values ($\dot{\bar{E}}$). The shaded area in each of these distributions, then, represents the probability that the momentary effective excitatory potential will on a particular trial be greater than the threshold value, L. But this probability value, which is designated as $p(\dot{\bar{E}} > L)$, also gives the probability of a response occurrence, for in the postulate concerning L it is assumed that a response occurs only when the momentary effective excitatory potential is greater than L. That is, R_p has the same value as $p(\dot{\bar{E}} > L)$. If now we were to plot the areas of the shaded portions of these distributions for successive values of $(\bar{E} - L)$ we would obtain a normal integral function. There is no need, then, to make a special postulate, as Hull did, for R_P is necessarily a normal integral function of $(\bar{E} - L)$ by virtue of the previous assumptions made concerning O and L.

Turning now to the latency measure, we can also derive its relation to momentary effective excitatory potential instead of making an arbitrary postulate as Hull did. The derivation is too lengthy and involved to give here. Suffice it to say that it involves assigning a value, *h*, to represent the average duration of a momentary O. We need then only to be able to calculate the mean expected number of successive momentary O values that will occur before one sufficiently small to produce a superthreshold momentary excitatory potential results. This value, in turn, may be shown to be a function of the probability of getting such an O value. The outcome of the derivation is shown in the following equation, in which t is the average latency per trial and h is the average duration of a momentary O:

$$t = \frac{h}{p(\dot{\bar{E}} > L)}$$

If we take the reciprocal of t we have the relation between a measure of speed of getting into action (S) and the momentary effective excitatory potential as follows:

$$S = \frac{1}{t} = \frac{p(\dot{\bar{E}} > L)}{h}$$

One of the interesting implications of this function is that if we now were to plot S against N, the number of trials, we should expect to get the same *shape* function that we would get with the probability of response measure. Figure 3 shows plots of this hypothetical probability value $[p(\dot{\bar{E}} > L)]$ as a function of N for a number of different values of the parameters that determine the level to which E will grow. Those of you who are familiar with curves of learning employing such a speed measure will recognize that they conform closely to these theoretical curves.

As in the case of response probability and latency it may also be shown that there was no need for Hull to make a special postulate concerning the relation between the third response measure—number of trials to extinction (R_N)—and the intervening theory. In this instance assumptions already made concerning the development and dissipation of inhibitory potential (I) lead to definite implications as to the relation between R_N and effective excitatory potential ($\bar{E}$). The relationship may be shown to depend upon whether massed or

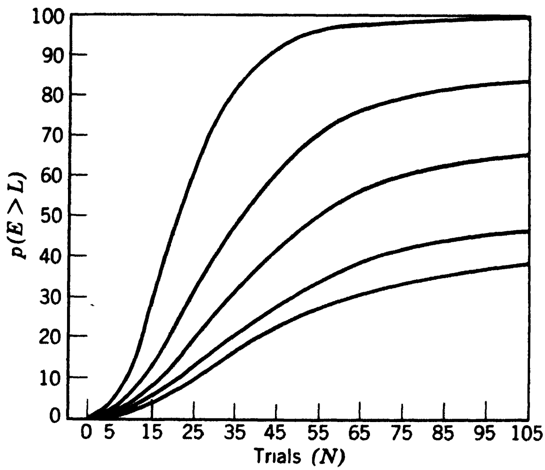

Fig. 3. Family of theoretically derived curves of the proportion of superthreshold momentary effective excitatory potentials [$p(\bar{E} > L)$] as a function of number of training trials for different growth curves of excitatory potential ($\bar{E}$).

distributed conditions of practice obtain. I shall not go further into the nature of the implied relations, as my primary purpose has been to bring out the point that it was unnecessary for Hull to make this final set of assumptions in the case of these three measures.

Although I have not had an opportunity to examine thoroughly the more recent theoretical formulations of Hull, including that presented in the first chapter of his new book, *A behavior system*, it is my impression that the same difficulties exist in them. Furthermore, Hull has made a number of changes in this portion of his theory, particularly in his conception of the oscillatory potential, O, which, in my opinion, are not for the best. For example, I would question the change from his original conception that the dispersion of O is invariable throughout learning to the conception that O begins with zero dispersion and increases as a growth function to a maximum dispersion in the first eight or nine trials. The reason that Hull gives for this change is that, once the response in instrumental learning has risen above the reaction threshold and has been reinforced, it relatively infrequently fails to occur during the subsequent learning. Just why Hull thought this fact to be contrary to his original conception is difficult to understand, for obviously in instrumental learning it is just a matter of waiting for the oscillatory *O* value to become sufficiently small to produce a momentary superthreshold excitatory potential. This happens in instrumental learning because the experimenter usually does wait for the response to occur. It does not

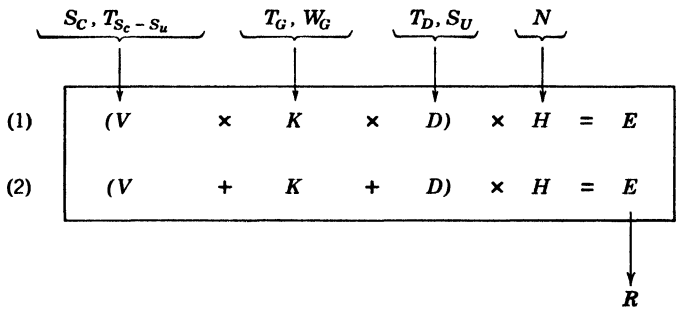

Fig. 4. Diagram showing relation between motivational (*V*, *K*, *D*) and learning variables (*H*).

happen in classical conditioning because the UCS does not wait for the conditioned response to occur. There are other reasons why I think the older conception is superior to the new, including some experimental evidence on the shape of frequency curves of classical conditioning, but time will not permit their discussion here.

I should like to direct your attention now to a second aspect of our theory in which there have been a number of new developments. Figure 4 presents that portion of our theoretical structure that deals with the motivational factors *V*, *K*, and *D* and their relations to the learning factor, *H*. In this diagram the experimentally manipulable variables are shown at the top of the figure, the intervening variables inside the rectangle, and the response variable at the bottom right of the figure. The arrows indicate which experimental variables are assumed to determine each of the intervening ones.

Equation 1 in the diagram represents the most recent Hull conception as to how the intervening motivational variables *V*, *K*, and *D* interact or combine with each other and with the learning factor *H* to determine *E* and hence response strength. Thus, on the basis of the available evidence in classical and instrumental conditioning, Hull assumed that each of the three motivational factors combined with *H* in a multiplicative fashion to determine *E*. Lacking evidence as to how *V*, *K*, and *D* combined with each other, Hull made the arbitrary working hypothesis that they also combined multiplicatively with each other.

Now I should like to call attention once more to the fact that Hull developed this theory in a purely *ad hoc* fashion to fit the known facts of these simple conditioning studies. Whether this motivational theory is adequate for more complex learning situations remains, of course, a question. Obviously, the manner in which such a question should be decided is to derive the implications of the theory for each complex learning situation and then check them against the experimental facts. Although this procedure seems obvious, as I have said, it has not always been followed. In its stead the flat assumption seems to have been made by some of our critics that an increase in drive strength *always* leads to a higher level of performance in more complex situations just as it does in simple learning. When this assumed result fails to occur, the theory is denounced and the futility of attempting to explain more complex behavior in terms of theoretical constructs from simpler situations is proclaimed.

The most charitable interpretation that can be made of this misuse of theory is that it is due to a failure to appreciate that the application of a theory to any particular situation involves not only a consideration of the laws or hypothetical relations postulated in the theory, but also what are referred to as the initial or boundary conditions of the situation. The logical implications of a theory are a joint function of the postulated relations or laws and the particular combination of conditions or variables operating in the situation. As I will now attempt to show by means of some examples, the implications of our theory as to the effect of varying drive strength in complex learning situations are not what has been so naïvely assumed. The implications differ, as we shall see, from one situation to another, and many other factors in addition to the manipulation of the strength of a single particular need must be taken into account.

One of the most important of the factors determining the effects of drive variation on performance in learning situations is the nature and complexity of the initial response hierarchy. If the learning task is a simple one in which there are no competing responses but only a single *S–R* tendency as in classical conditioning, or a set of more or less isolated single *S–R* tendencies as may be arranged in paired-associate learning, a high drive level should be expected to lead to a relatively high level of responding. Thus in classical conditioning there is, by virtue of the control of the stimulus conditions, a single highly dominant response. As training proceeds, the CS acquires habit strength (H) for this response. The excitatory strength (E) of the CS to evoke the response depends, it is assumed, on the product of this H and the level of D, i.e., $E = f(H \times D)$. The higher the value of D, the greater E will be and hence the greater will be the response strength. In more complex learning situations involving a hierarchy of competing responses, however, the effect of drive-strength variation will depend upon a number of factors the influence of which must be evaluated in each specific situation.

Ideally, the instrumental learning situation is also one in which there is but a single S–R tendency, and if this were so the same prediction could unqualifiedly be made for it as for classical conditioning. In actual practice, however, instrumental learning situations are limiting cases of trial and error learning in which there is an initial hierarchy of S–R tendencies. Although the experimenter usually chooses the strongest response in the hierarchy to reward, or attempts by various means to make it the strongest, the presence of the other competing responses needs to be considered. I shall not attempt to treat this situation in detail here, as it is much more complicated than has been realized. It is perhaps sufficient to state that after considerable training the habit strength of the reinforced response is so much greater than that of the competing responses that the latter have little or no effect. At this stage it may be predicted that the response speed would be directly related to drive level, as in classical conditioning.

Let us now consider a selective learning situation in which there are two competing responses of about equal strength at the beginning of training. A very simple situation of this type is one in which we have two alternative responses both of which lead to a reward, but one of which involves a longer delay of the reward than the other (e.g., pressing one of two bars, entering a left or a right alley). We recently employed such a situation in an experiment to test an alternative hypothesis to that of Hull's as to how the motivational factors K and D combine. This hypothesis is shown in the second equation of Fig. 4. It will be seen that the hypothesis assumes K and D combine in some additive manner rather than multiplicatively as assumed arbitrarily by Hull. We were interested in comparing the performance of two groups of white rats under different hunger levels in such a situation where the number of responses was kept equal by means of forced trials. One of the responses involved a delay in reward of 1 second, the other a delay of 5 seconds. The deprivation times were 3 hours and 22 hours. The prediction for the additive assumption is that there will be no difference in the per cent choice of the short-delay response over the long-delay under the two levels of drive strength. The derivation may be shown algebraically as follows:

$$E_S = H(D + K_S)$$

$$E_L = H(D + K_L)$$

$$E_S - E_L = H(D + K_S) - H(D + K_L)$$

$$= H(K_S - K_L)$$

Thus it will be seen that the difference between the excitatory potentials of the short- and long-delay responses and hence the percentage choice of one over the other will be a function of H (i.e., the number of reinforced trials on each response) and K (i.e., time of delay of reward) but not a function of D (i.e., time of deprivation).

The implication of Hull's multiplicative assumption for the same experimental situation, on the other hand, may be shown to be just the opposite, i.e., that performance will be a function of D. The outcome of the experiment, although not entirely clear cut, tended to favor the additive assumption, for a significant difference between the two drive groups failed to appear. My own expectation is that appropriate experimentation will demonstrate that all three motivational factors will be found to combine with each other in an additive manner as suggested in the second equation of Fig. 4.

Let us turn now to another competing-response situation—one that again involves two alternative responses in the hierarchy, such as a left and a right alley in a simple T-maze. In this situation one of the responses leads to reward; the other to an empty goal box *with no opportunity to correct.* Assuming the responses are equal in strength to begin with, or that the correct one has somewhat greater habit strength than the wrong one, it may be predicted from our theory (either Hull's or my version) that a higher drive level will lead to a higher level of performance. The derivation is as follows:

$$E_+ = H_+ \times D$$

$$E_- = H_- \times D$$

$$E_+ - E_- = (H_+ \times D) - (H_- \times D)$$
$$= D(H_+ - H_-)$$

Here it will be seen that a difference in level of D does determine the magnitude of the difference in the excitatory potentials of the two competing responses and hence the probability of choice of one over the other. Quite in contrast to the previous situation, then, the additive hypothesis does imply a difference in performance level under different drive levels in this non-correction situation.

Now, as some of you may know, there have been a number of experimental studies of the effect of drive level on performance in this type of selective learning situation, and the findings have been highly conflicting. Obviously there must still be some other factor operative in these experiments, differential variation of which has produced these conflicting findings. What could such a factor be? One possibility that is suggested by our theory has to do with the relative habit strengths of the two competing responses at the start of the experiment. You will recall that one of the conditions involved in our prediction was that, initially, either the responses were equal in habit strength or the correct one was slightly stronger. This is an extremely important condition, for the implication in which the incorrect response is initially stronger in habit strength than the correct (reinforced) response is just the opposite—namely, that the high-drive group would start out at a relatively lower level of correct choice and hence would make more errors than the low-drive group.

If now the experimenter should throw together into the same group subjects with opposite initial response preferences, we would have two opposing effects that would tend to cancel each other and thus result in no difference in performance under the different drive levels. If the experimenter should have only subjects who have no initial preferences or only slight preference for the correct response, a difference under varying drive levels in favor of the high-drive group would be obtained. Should the experimenter have only subjects with strong initial preferences and should he train them against their preference, the low-drive group should perform the better, at least in the earlier stages of the learning.

The predictions just made, it should be noted, were in terms of the per cent of choice of one response over the other. So far as this frequency or probability measure is concerned, we can ignore other possible superthreshold response tendencies that undoubtedly are present. We simply confine our scoring to noting the occurrence of one or the other of these two selected responses. But the situation is very different in the case of a time measure, e.g., time required to make either response. Particularly would this be so if there were other response tendencies in the hierarchy (such as trying to climb out of the apparatus) that were stronger than those of approaching and entering either of the alleys. In this event it may be shown that such a time measure should be longer for high drive levels than for low drive levels. That is, the animals would be expected to take longer to make either of the responses under the high-drive conditions than under the low-drive conditions.

I hope that these considerations have given you a better appreciation of some of the problems that are involved in making such predictions from theory and that they will convince you of the importance of the point I made earlier: namely, that the logical implications of a theory are always a joint function of the postulated relations or laws and the so-called initial conditions.

Let me finish by just mentioning briefly one or two further considerations. The matter of the type of measure used is of particular importance for the problem-box type of situation in which the measure most usually employed is the time taken to make the correct response. Such learning situations typically involve a large number of responses in the initial response hierarchy, with the correct goal-attaining one being relatively low in habit strength. Depending on a number of factors, such as the relative position of the correct response in the hierarchy, the magnitude of the initial habit-strength differences among the alternative responses, and the number of responses of fairly weak habit strength, different predictions as to the effects of varying drive strengths would be made. Since these initial values are usually unknown it is impossible to use them in making precise tests of our theory. As a consequence we have been forced to design new complex learning situations in which the initial response hierarchies can be determined in terms of experiences prior to or during the experiment.

An interesting variant of this latter procedure is the series of motivational experiments we have been conducting at Iowa over the past few years. Employing human subjects, these studies have attempted to manipulate drive level by selecting two groups in terms of their extreme scores on a test of manifest anxiety. Those of you who are familiar with these experiments will recall that we found that subjects at the high (anxiety) end of the scale exhibited a significantly higher level of performance in classical conditioning than did subjects from the low (non-anxiety) end of the scale [*20, 21, 22*]. Interpreting the high end of the scale as reflecting high drive level and the low end as reflecting low drive level, it was then predicted that the opposite finding, i.e., superior performance on the part of non-anxious (low-drive-level) subjects, would be found in more complex learning situations in which there was a high incidence of strong, incorrect, competing responses. This prediction has been confirmed in three separate experiments involving rote serial learning [*13*], a verbal maze [*23*], and a stylus maze [*4*]. The stylus-maze experiment is of particular interest because of the fact that the same subjects were run in an eyelid conditioning experiment. Whereas the anxious subjects were superior in the conditioning situation, the opposite was the case in the maze experiment.

We have just recently extended this research to paired-associate learning. We constructed one list of words in which every effort was made to make the correct *S–R* tendencies relatively strong and to minimize the possibility of competing response tendencies. For such a list it was predicted that anxious (high-drive) subjects would perform better (make fewer errors) than non-anxious (low-drive) subjects. In a second list an attempt was made to maximize the number and strengths of the competing responses and thus obtain a list on which the non-anxious subjects would be superior to the anxious subjects. The procedure involved the manipulation of the synonymity of the stimulus words and the strength of associative connections of the paired words. Preliminary results for only nine subjects in each group are shown in Table 1.

It will be seen that, just as was predicted, the anxious subjects performed better on list 1, in which the number and strength of competing responses were minimized, whereas they were inferior on list 2, in which competing responses were stronger than the correct re-

TABLE 1 *Mean number of errors made by anxious and non-anxious subjects on list 1 (no or low competition) and list 2 (high competition)*

Subjects	List 1	List 2
Anxious (N = 9)	15.2	98.7
Non-anxious (N = 9)	27.4	79.4
p	<0.01	<0.05

sponses. It should be realized that different subjects were involved in the two lists, but a comparison of the groups on a preliminary practice list showed them to be comparable.

In concluding I cannot forego pointing out that the type of additive hypothesis that I have been employing—a curvilinear one along the lines of the manner in which Hull assumed habit strengths to summate —requires that consideration also be given to the strengths of the *V* and *K* values when making predictions as to what the effects of variation in *D* will be. Very briefly, the predictions we have made hold when the values of *V* and *K* are relatively low. If these values are maximized there will be a general wiping out of the influence of variations of *D*, and performance will tend to be quite independent of them. This latter statement holds, of course, even within the factor *D* itself. Thus, as has been shown in a number of instances, different strengths of the hunger need do not produce performance differences if a strong pain or anxiety condition is also present.

Significantly enough, it is the global-minded person who, while giving full lip service to the principle of taking all the factors into account, invariably fails to do so when disproving one of our theoretical predictions. This seeming perverseness is only apparent, however; for, lacking any type of theoretical analysis, such a person literally doesn't recognize the potential factors and their possible action. Although there are dangers in the biases of a theorist, this danger is negligible if a genuine effort is exerted to make the hypothesized relations as specific and precise as possible. Theoretical biases thrive only in the vague types of formulation that fail to meet this specification.

REFERENCES

1. Bush, R. R., and Mosteller, F., A mathematical model for simple learning, *Psychol. Rev.*, 1951, *58*, 313–323.
2. Dollard, John, and Miller, N. E., *Personality and psychotherapy*, New Haven, Yale University Press, 1952.
3. Estes, W. K., Toward a statistical theory of learning, *Psychol. Rev.*, 1950, *57*, 94–107.
4. Farber, I. E., and Spence, K. W., Complex learning and conditioning as a function of anxiety, *J. exp. Psychol.*, 1953, *45*, 120–126.
5. Grice, G. R., An experimental study of the gradient of reinforcement in maze learning, *J. exp. Psychol.*, 1942, *30*, 475–489.
6. Gulliksen, Harold, A rational equation of the learning curve based on Thorndike's law of effect, *J. gen. Psychol.*, 1934, *11*, 395–431.
7. Guthrie, E. R., *The psychology of learning*, New York, Harper and Brothers, 1935.
8. Householder, A. S., and Landahl, H. D., Mathematical biophysics of the central nervous system, *Mathematical Biophysics Monogr.*, Series 1, Bloomington, Ind., The Principia Press, 1945.
9. Hull, C. L., *Principles of behavior*, New York, Appleton-Century, 1943.

10. Hull, C. L., *Essentials of behavior,* New Haven, Yale University Press, 1950.
11. Hull, C. L., *A behavior system,* New Haven, Yale University Press, 1952.
12. Miller, Neal E., and Dollard, John, *Social learning and imitation,* New Haven, Yale University Press, 1941.
13. Montague, E. K., The role of anxiety in serial rote learning, *J. exp. Psychol.,* 1953, *45,* 91–97.
14. Perin, C. T., Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction, *J. exp. Psychol.,* 1942, *30,* 93–113.
15. Pitts, W., A general theory of learning and conditioning, *Psychometrika,* 1943, *8,* 1–18, 131–140.
16. Rashevsky, N., *Mathematical biophysics,* Chicago, University of Chicago Press, 1938.
17. Spence, K. W., The postulates and methods of "behaviorism," *Psychol. Rev.,* 1948, *55,* 67–78.
18. Spence, K. W., Mathematical formulations of learning phenomena, *Psychol. Rev.,* 1952, *59,* 152–160.
19. Spence, K. W., The nature of discrimination learning in animals, *Psychol. Rev.,* 1936, *43,* 427–449.
20. Spence, K. W., and Taylor, Janet A., Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning, *J. exp. Psychol.,* 1951, *42,* 183–188.
21. Spence, K. W., and Farber, I. E., Conditioning and extinction as a function of anxiety, *J. exp. Psychol.,* 1953, *45,* 116–120.
22. Taylor, Janet A., The relationship of anxiety to the conditioned eyelid response, *J. exp. Psychol.,* 1951, *41,* 81–92.
23. Taylor, Janet A., and Spence, K. W., The relationship of anxiety level to performance in serial learning, *J. exp. Psychol.,* 1952, *44,* 61–64.
24. Thompson, M., Learning as a function of the absolute and relative amounts of work, *J. exp. Psychol.,* 1944, *34,* 506–515.
25. Thurstone, L. L., The learning function, *J. gen. Psychol.,* 1930, *3,* 469–493.
26. Tolman, E. C., Operational behaviorism and current trends in psychology, *Proc. 25th Anniv. Celebration Inaug. Grad. Stud.,* Los Angeles, The University of Southern California, 1936, pp. 89–103.
27. Tolman, E. C., The determiners of behavior at a choice point, *Psychol. Rev.,* 1938, *45,* 1–41.
28. Tolman, E. C., "Discussion," *J. Pers.,* 1949, *18,* 48–50.

Stimulus-Response Theory as Applied to Perception

DELOS D. WICKENS

INTRODUCTION

I should like in this paper to consider a type of perceptual activity which seems to be of major concern to many practicing clinical psychologists but which as yet has not been systematically handled within the concepts of stimulus-response psychology. I am referring to the fact that, given a particular complex environmental situation, the client will react to a certain aspect of the situation and disregard other aspects. Out of all the events that have occurred in some series of social interactions he may be alert only to those which imply some criticism of himself; he may even twist events, warping them so that they fit readily into this inaccurate perceptual schema. Given such selected and distorted perceptual data, he is in a sense logically justified in reacting, let us say, with antagonism to his fellow beings. As a result many a clinician has concluded that the key to maladjustment lies in perception, and that if only the client could be made to perceive his environment in a more accurate manner, his maladjustive responses would be eliminated. In a sense this clinician seems to feel that certain kinds of responses are appropriate to certain kinds of perceptions, whereas other responses are appropriate to other perceptions; change perception and you change behavior.

When the clinician thinks in this fashion he implicitly or explicitly assumes that there is no necessary and invariant connection between a certain physical environment and a certain perception. He is drawing a distinction between the physical environment and the behavioral environment [8]. Such distinctions seem necessary.

Now it is certainly true of early stimulus-response theory that there was little if any deference paid to the possibility of variance between physical stimulus and some perceptual activity in the organism. Indeed, the word "perception" almost seems not to exist in the vocabulary of this early theory. To a certain degree it was admitted into

Hull's system [6, 7] with his differentiation of the capital S and the lower case s, the capital referring to the physical stimulus, the lower case to the resultant neural activity in the organism. Hull's concept of afferent neural interaction, $\check{S}$, goes even farther. When we recall that his H_R, his I_R, and his E_R have as their prefixes the lower case s, we can see that at least to a limited degree this more recent formulation of S–R theory contained a perceptual term. I think Guthrie had always included such a term in his concept of the movement-produced stimuli to which responses become attached [4].

Even though there are these provisions for perceptual data in S–R theories, the research which has been done within the framework of these types of theory has usually disregarded perceptual problems. I do not mean to imply that this disregard vitiates much of the experimental work, for actually what these experimenters have done for the most part is to control the factors which might lead to perceptual variance. These controls have been achieved through the selection of certain restricted types of situations in which to perform their experiments, through pretraining, and through the directions given to the subjects. As a result these researchers have, I think, been able to operate as if the stimulus were the physical event. My biases lead me to feel that such work has led to the identification of a number of important and basic principles of behavior.

Actually my purpose in this paper is to attempt to employ some of these same principles to predict certain perceptual phenomena. Essentially I am taking the position that these perceptual characteristics with which our hypothetical clinician is concerned are molar phenomena which can be predicted from certain of the molecular S–R postulates. Needless to say, I shall hereafter attempt to make my task an easier one by forgetting about the hypothetical clinician and client, and choose my examples from simpler experimental situations. They are the type of experimental situations which are described as offering evidence for the operation of perceptual sets or perceptual biases. These sets and biases are indicated in these experiments by the tendency of the subject to respond positively to certain characteristics of a complex stimulus object and to disregard other aspects of it; in other words, to behave in the laboratory situation as the unadjusted client behaves in his everyday life. Our experimental examples have a real advantage over the clinical cases in that the experimenter, by manipulating his experimental conditions, throws some light on the conditions necessary for such perceptual biases. Such was true of the studies by

Lawrence [9, 10] and a later experiment by Eckstrand [1] in which human beings rather than rats were the subjects.

The behavior under consideration will be illustrated by an analysis of an experimental situation similar to the one used by Eckstrand [1]. Subjects are presented singly with stimuli to which they are to react by pressing one of three keys. For each stimulus one and only one key or response is correct. Actually there are nine different stimuli, but they are made by combining three forms with three colors. The solution to the problem consists of responding only to the color aspect of the stimulus regardless of form. The experimental design is illustrated in Table 1. We will assume that the colors employed are

TABLE 1 *Design of the stimulus-response relationships in the hypothetical experiment used to develop a perceptual bias to respond to the dimension of color in a transfer situation*

Relevant Stimuli	Irrelevant Stimuli F_1	F_2	F_3	Correct Response
C_1	Yellow triangle	Yellow circle	Yellow square	R_a
C_2	Green triangle	Green circle	Green square	R_b
C_3	Blue triangle	Blue circle	Blue square	R_c

yellow, green, and blue, and that the forms are triangle, circle, and square. To the yellow stimulus the only response that is reinforced is R_a, regardless of the form with which it is associated; any other response is not reinforced. The green stimulus is reinforced only if the response R_b is made, and the blue only if R_c is made. Hereafter these colors will be referred to by the letter C and appropriate subscript, and the forms by the letter F and appropriate subscript.

The results of the experiment by Eckstrand suggest that, after subjects have solved this problem and are given a new problem in which different colors and different forms are employed, they will learn it with greater ease than if the first problem had not been presented and solved [1]. One may say that these subjects have learned to pay attention to color, have developed a concept that color is the key to the solution of the problem, or are perceiving selectively. The question is whether it is possible to predict this behavior from molecular S–R concepts.

The following assumptions will be made in applying S–R theory to this area of behavior:

1. If reinforcement occurs following a response to a stimulus, the tendency to make that response to that stimulus is strengthened. The

strengthening is non-selective in nature, and it occurs for those aspects of the total stimulus which are arbitrarily wrong as well as those which are arbitrarily right.

2. Conversely, if the response is not followed by reinforcement, the tendency to make that response to that stimulus will decrease.

3. Stimulus generalization will occur under both the strengthening and the weakening conditions. The form of the gradient I will employ is one which is bell-shaped rather than concave upward as is Hovland's [5]. This gradient is based upon some recent research conducted at Ohio State by Wickens, Schroeder, and Snide. I am also using the same form and shape of the generalization curve for excitation and inhibition. This may be incorrect. Actually the exact form of the generalization gradients is not important at this stage of the analysis, and if later research should indicate that such gradients do not hold generally, no major modification would need to be made.

4. The increment in habit strength gained from one reinforcement is equal to the decrement resulting from one non-reinforcement. This assumption is not crucial, and it is made only for purposes of simplifying the exposition. Exact empirical data concerning this assumption are lacking.

5. The learning is continuous in nature and not dependent upon the subject's hypothesis [11]. Indeed, the central position taken in this paper is that hypotheses are not the cause of the learning, but are in the nature of the subject's verbal expression of the habits which have been acquired.

THE CONDITION OF EQUAL ASSOCIATION OF PARTICULAR RELEVANT AND IRRELEVANT STIMULI

Using these assumptions, we will now consider the strength of the various habits the subject has acquired after twenty-seven stimulus-response events. We will assume that during these twenty-seven events each of the nine stimuli has been presented three times, and each of the three responses has been made once to each of the nine stimuli. There is nothing essential about this sequence of events, and the arguments presented would hold for another pattern of responses as well; it is arbitrarily chosen for purposes of simplifying the exposition.

Table 2 summarizes the results of this procedure in terms of the occurrence of reinforcement and non-reinforcement associated with each stimulus-response relationship. It will be noted that reinforcement was given for the response R_a only when it was made in the presence of the stimulus C_1. At all other times this response occurred, it did not receive reinforcement. A similar state of affairs holds for C_2 and R_b and for C_3 and R_c.

TABLE 2 *The pattern of reinforcement and non-reinforcement associated with each stimulus and each response following 27 trials*

Stimuli		Responses		
Color	Form	R_a	R_b	R_c
C_1	F_1	+	−	−
	F_2	+	−	−
	F_3	+	−	−
C_2	F_1	−	+	−
	F_2	−	+	−
	F_3	−	+	−
C_3	F_1	−	−	+
	F_2	−	−	+
	F_3	−	−	+

We will now consider what effect this sequence of events has on the tendency to evoke the R_a response *only* over the entire dimension of color. Three reinforcements of this response were given at point C_1

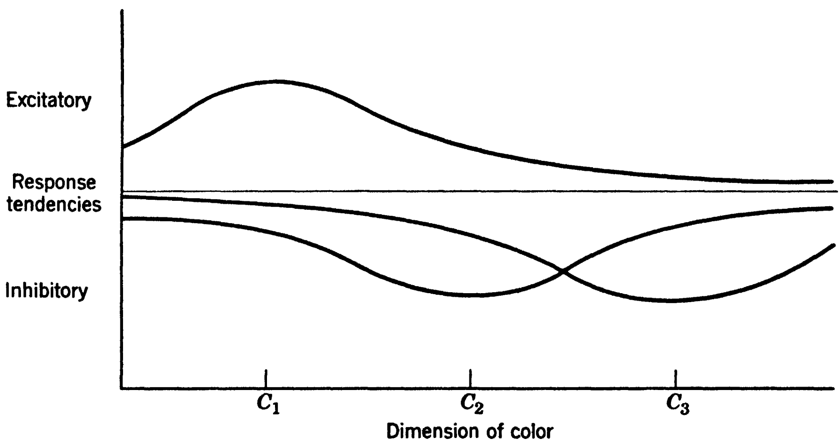

Fig. 1. The separate inhibitory and excitatory tendencies to make the response R_a across the dimension of color.

on the continuum, thus producing a generalization gradient of excitatory tendencies across the entire continuum. This is the gradient shown above the zero line in Fig. 1. However, three non-reinforcements were given at points C_2 and C_3, and these generate the two curves below the zero line.* If these three gradients are summated algebraically, the resultant tendencies to make R_a across the entire dimension of color are shown in Fig. 2. It is apparent that there is a tendency to make this response to stimuli in a range close to C_1, but to avoid making this response in the ranges surrounding C_2 and C_3.

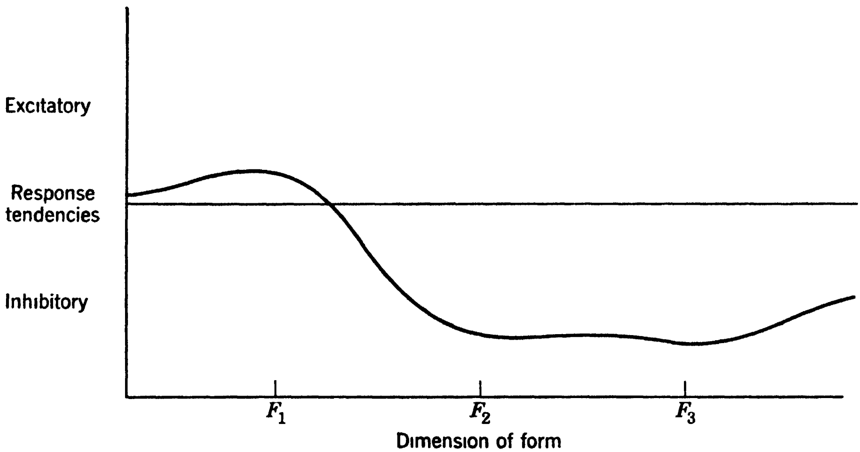

Fig. 2. The algebraic sum of the inhibitory and excitatory tendencies to make the response R_a across the dimension of color.

A similar result will occur for C_2 and its response R_b and for C_3 and its response R_c. When these three separate S–R connections are placed upon the same continuum, the curves shown in Fig. 3 are generated. The curves indicate that there is a tendency for one or the other responses to be evoked by new stimuli along most of the C dimension. Thus, if a group of new colors were employed, these colors would tend to evoke this class of responses.

It will be noted that there are gaps in the excitatory curves in various portions of the continuum. These result in part because of the relatively high frequency of wrong or unreinforced trials to correct or reinforced trials. As we will point out later, these gaps will ordinarily be eliminated with increased training and a consequent increased proportion of correct responses. Also, these gaps would

* It is apparent from this and other figures that we have represented the stimuli as being equidistant from each other along the dimensions of both form and color. The arguments presented in the text are not, however, restricted to such a state of affairs.

have been less extensive if the generalization curves for inhibition were steeper than those for excitation.

Turning now to the effects of the training upon the *F* stimuli, we find that the results are quite different in nature. As shown in Table 2 every stimulus has been given one reinforcement to every response, but it has also been given two non-reinforced trials for every response. In other words, the inhibitory tendencies are at every point more frequent than the excitatory tendencies at that point. The net result

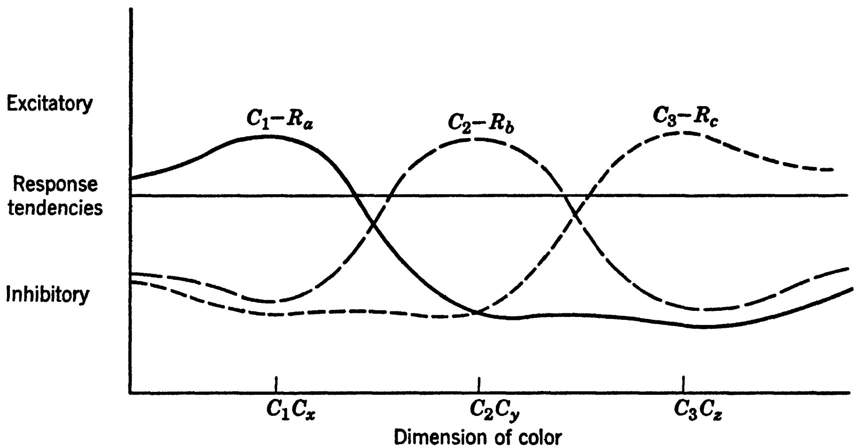

Fig. 3. Response tendencies for all color stimuli across the color dimension.

of this treatment is inhibition across the entire continuum, and the curve shown in Fig. 4 is generated. This curve will be the same for all of the *F* stimuli.

The structure of this training has therefore developed habit tendencies within the subject such that, if he is presented with a new problem using new colors and new forms, responses are not likely to be evoked by the aspect of form, but they are likely to be evoked by the aspect of color. This kind of behavior has the characteristics of a perceptual set to respond to color. It is a form of selective perception.

THE EFFECT OF NUMBER OF TRIALS

As the subject begins to respond correctly in this situation, and reinforcements pile up to the exclusion of non-reinforcements, a change will occur in the nature of the response tendencies to both the *C* and the *F* stimuli.

One effect the increasing proportions of correct responses will have upon the tendencies to respond to the relevant or C stimulus is to decrease the inhibition associated with any particular C stimulus and the two wrong responses. In our example considerable inhibition has been built up for the S–R connections of C_1–R_b and C_1–R_c. As these erroneous responses are dropped out because of the increasing

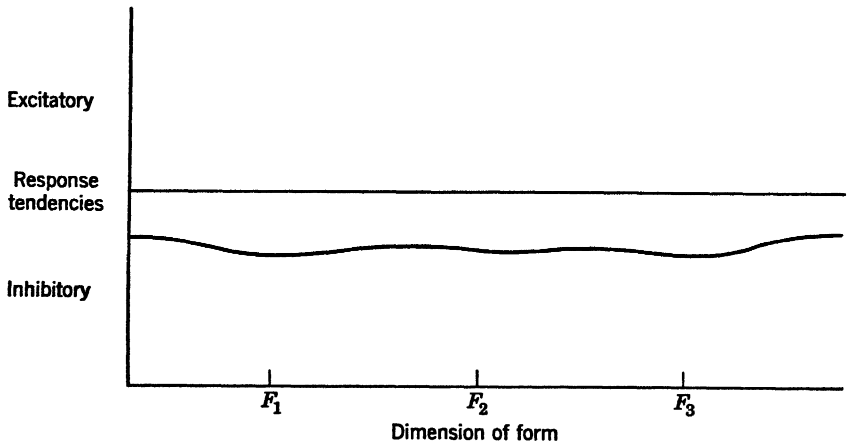

Fig. 4. The algebraic sum of the inhibitory and excitatory tendencies to make any response across the dimension of form.

strength of the C_1–R_a connection, less and less inhibition will accrue to them.

This effect will be even more marked in the instance of the F or irrelevant stimuli. These stimuli ride along with the relevant ones, and each time a correct response is made to the relevant stimulus the associated irrelevant stimulus profits from the consequent reinforcement. Gradually the inhibitory potentials associated with the irrelevant stimuli will be wiped out, and they will be replaced by excitatory potentials. The patterns of distribution of the excitatory potentials associated with these stimuli across the dimension will differ markedly from that characteristic of the relevant stimuli. The distribution for the relevant stimuli is asymmetrical, with particular stimuli being positively associated with certain responses and negatively associated with others. Since, however, each irrelevant stimulus is associated equally often with each relevant stimulus, and thereby reinforced at each response point, the distribution of tendencies will be symmetrical. Each F stimulus will develop an equal tendency to produce each response.

In summary, as N increases (and in the ordinary course of events the number of correct or reinforced responses increases), the irrelevant stimuli will begin to acquire excitatory tendencies. Since, however, particular relevant S–R connections already have a head start, we should not expect that subjects would eventually fall into error and begin responding to the irrelevant or F stimuli. There is evidence in an experiment by Grant and Berg [3] that increased training makes it easier for the subject to use as relevant a dimension that has previously been irrelevant.

THE EFFECT OF RESPONSE GENERALIZATION

According to the concept of response generalization, when a particular S–R connection is reinforced, the tendency to give similar responses to the stimulus is also strengthened. Although this concept suffers in usefulness because of the difficulty of rating responses on their degree of similarity, it has some empirical support [*12*, *13*]. In the present hypothetical experiment, one would seem justified in assuming that the pressing of one key is highly similar to the pressing of another key. If such is the case, this mechanism would, as the subject begins to respond correctly, serve to decrease even more the inhibition originally built up for particular relevant stimuli and the wrong responses. Thus, when to C_1 the response R_a is made and reinforced, excitatory potential spreads to R_b and R_c. The magnitude of increment would not be as great for these two responses as for R_a, however. This mechanism would so raise the excitatory level that the gaps of inhibition shown between C_1 and C_2 and also between C_2 and C_3 in Fig. 3 would no longer exist. Thus response generalization increases the tendency to make key-pressing responses across the entire dimension of color. It would also tend to eliminate some of the asymmetry of the pattern of response potential associated with the C stimuli. Since the F stimuli are already symmetrical, the mechanism would simply add a constant to the distribution of inhibitory potential without changing the symmetry of the distribution.

THE EFFECT OF STIMULUS-RESPONSE RELATIONSHIP ON THE FIRST AND SECOND TASK

The magnitude of transfer from one task to a new one will be, according to this analysis, a function of the location of the new stimuli on the dimension and the responses that are required to be made to them [2]. In a new task a problem may be set for the subject in which the relevant cues are drawn from the same dimension as those employed in the first task. These stimuli may be located at points C_x, C_y, and C_z in Fig. 3. If the response R_a is to be given to C_x, R_b to C_y, and R_c to C_z, then a considerable amount of positive transfer will occur. If, however, the responses required are not consonant with the original learning (for example, R_a to C_z, R_b to C_x, and R_c to C_y), negative transfer or less positive transfer will occur [2]. This latter hedging statement arises from several considerations. One of these considerations is response generalization which would heighten the tendency to make any response to any of the relevant stimuli. Another is that human subjects may verbalize, saying perhaps, "Its color" rather than "Yellow for the middle key." If the verbalization is correct, this mediating response that is based on previous learning would be reinforced, strengthening the entire dimension of color.

In general it will be noted that this analysis does not completely free perception of the overt responses which have become attached to the various stimuli on the dimension. It implies that perceptual transfer would not be the same in experiments where the first and second task utilized the same responses as in experiments wherein the responses on the two tasks differed. Response generalization and language mediation would, however, lessen such differences.

THE EFFECT OF UNEQUAL ASSOCIATION OF PARTICULAR RELEVANT AND IRRELEVANT STIMULI

Under actual life conditions it is very likely that the complete symmetry of association of the irrelevant cues with the relevant cues may not occur. In fact it is probable that in the usual life situation symmetry of the sort that can be attained in laboratory experiments is the exception rather than the rule. More often than not, the natural situation is one in which one particular irrelevant stimulus is more frequently associated with one particular relevant stimulus than with any of the others. Thus the experimental design previously described might be modified so that F_1 was associated with C_1 60 per cent of the time that C_1 occurred, and was associated with each of the other stimuli 20 per cent of the time. F_2 could be linked to C_2, and F_3 to C_3, in a similar fashion. If this were done, then each F stimulus would receive a greater proportion of reinforcements and a smaller proportion of non-reinforcements for one particular response than for the other two. Each of the F stimuli would require relatively stronger tendencies to elicit one particular response than to elicit either of the other two responses.

The same procedure may be employed for this type of situation as that described in the first section, the procedure in which each set of color stimuli is associated with each response an equal number of times. In the present case, to achieve proper counterbalancing each C stimulus will be associated with each response five times and will

TABLE 3 *The frequency of reinforcement or non-reinforcement associated with each irrelevant stimulus in the 3–1–1 design*

Stimuli Presented		Responses Made					
		R_a		R_b		R_c	
Color	Form	+	−	+	−	+	−
C_1 (15)	F_1 (9)	3			3		3
	F_2 (3)	1			1		1
	F_3 (3)	1			1		1
C_2 (15)	F_1 (3)		1	1			1
	F_2 (9)		3	3			3
	F_3 (3)		1	1			1
C_3 (15)	F_1 (3)		1		1	1	
	F_2 (3)		1		1	1	
	F_3 (9)		3		3	3	

be reinforced or non-reinforced according to whether or not that response is correct for that stimulus.

The results for the present arrangement of relevant and irrelevant stimuli in terms of frequency of reinforcement and non-reinforcement for each F stimulus are given in Table 3. In this table the first major column denotes the stimuli presented; this column is subdivided into the color and form aspects of the stimulus. The numbers in parentheses indicate the frequency of presentation for each aspect. Thus C_1 is presented a total of fifteen times, nine times with F_1 and three times each with F_2 and F_3. The next group of columns designates what responses are made and whether or not they are reinforced. Note that R_a is always reinforced if made in the presence of C_1, but responses R_c and R_b are never reinforced if C_1 is present.

I shall not be concerned with the relevant stimuli in this design, but only the irrelevant ones, and the stimulus F_1 in particular. Going down the first major column, we see that F_1 receives three reinforcements in conjunction with R_a (these when it is associated with C_1) and two non-reinforcements (these when it is associated with C_2 and C_3). The predominance of reinforcements arises, of course, from its linkage with the relevant stimulus C_1. The result is greater excitatory than inhibitory strength for the S–R connection F_1–R_a. In the second major column, under R_b, however, there are a total of four non-reinforcements to one reinforcement. The same holds true for R_c. The resultant gradients are illustrated in Fig. 5, and they are summated algebraically in Fig. 6.

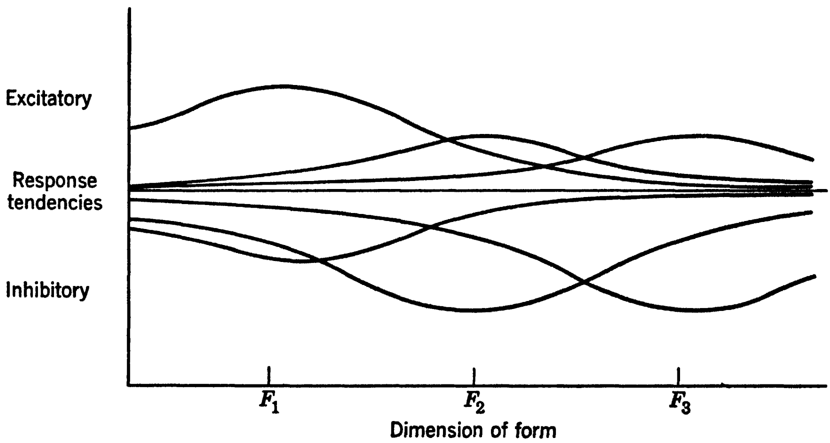

Fig. 5. The separate inhibitory and excitatory tendencies to make the response R_a across the dimension of form in the asymmetrical distribution of irrelevant cues.

A similar treatment for F_2 and R_b, and for F_3 and R_c, would also result in this asymmetrical patterning, R_b having its excitatory peak in conjunction with F_2, and R_c in conjunction with F_3. It is apparent that the over-all curves to be drawn for these F stimuli would be similar to the curves in Fig. 3. They would differ, however, in that the magnitude of excitatory potential would not be as high as in the former curve.

As the proportion of times is increased that any given irrelevant stimulus is associated with a particular relevant stimulus, the amount of excitatory potential acquired by the irrelevant stimuli increases. Thus, if the proportion had been 6–1–1 rather than 3–1–1 in the above example, the asymmetry of the habit loadings of the irrelevant stimuli would be even more marked and would become even more similar to the loadings for the relevant stimuli.

The present analysis implies that there is a basic inadequacy in the perceptual type of interpretation which classifies stimuli as either relevant or irrelevant. This *S–R* interpretation leads to the assumption of a continuity between the poles of the two extremes which for convenience will be called relevant or irrelevant.

I should like to mention an experiment that Mr. Harold Babb is doing at Ohio State on the effect of varying degrees of association of an irrelevant stimulus with the relevant one. The experiment is

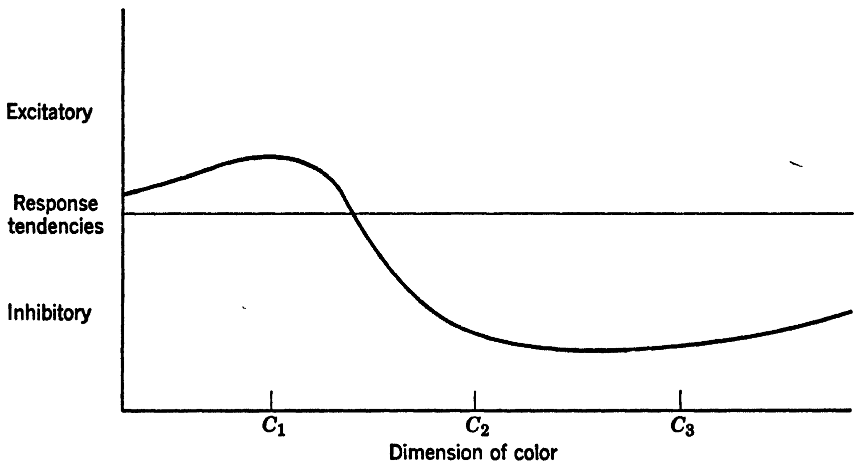

Fig. 6. The algebraic sum of the inhibitory and excitatory tendencies to make the response R_a across the dimension of form in the asymmetrical distribution of irrelevant cues.

modeled after those done by Lawrence [*9*, *10*]. Rats are the subjects, and their task is to learn a discrimination. During the first part of the experiment the discriminative cue is the presence or absence of some chains hanging in the alley. For the control group the alleys are gray, but for three experimental groups they are black or white. For one of these groups white is associated with chains, the correct cue, 30 per cent of the time; for another, 50 per cent—i.e., randomly; and for the third, 70 per cent. In the second part of the experiment the animals learn a black-white discrimination with white positive. The experiment is as yet incomplete, but the trend is for the 30 and 50 per cent groups to have the greatest difficulty in learning the black-white discrimination, the control group to rank next, and the 70 per cent to show the greatest ease. Thus, at least so far, the results conform with theoretical predictions.

SUMMARY

There is a considerable amount of evidence which indicates that a one-to-one relationship does not exist between the physical characteristics of a complex stimulus and its stimulating value for the organism. Instead the stimulating value seems to depend upon perceptual characteristics of the organism. This paper is an attempt to view the perceptual responses as mediating responses to which the overt responses are made, but which are predictable from a knowledge of the prior experiences of the organism. It is further assumed that the molecular postulates of *S–R* psychology, primarily the postulates dealing with reinforcement, non-reinforcement, and stimulus generalization, may be employed in making these predictions.

REFERENCES

1. Eckstrand, G. A., Cue attention habits as a factor in training, *A. F. Technical Report* 6566, August 1951.
2. Gagne, R. M., Baker, K. E., and Foster, H., On the relation between similarity and transfer of training in the learning of discriminative motor tasks, *Psychol. Rev.*, 1950, *57*, 67–79.
3. Grant, D. A., and Berg, E., A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem, *J. exp. Psychol.*, 1948, *38*, 404–411.
4. Guthrie, E. R., *The psychology of learning*, New York, Harper and Brothers, 1951.
5. Hovland, C. I., The generalization of conditioned responses: I The sensory generalization of conditioned responses with varying frequencies of tone, *J. gen. Psychol.*, 1937, *17*, 125–148.
6. Hull, C. L., *Principles of behavior*, New York, Appleton-Century, 1943.
7. Hull, C. L., *Essentials of behavior*, New Haven, Yale University Press, 1951.
8. Koffka, K., *Principles of gestalt psychology*, New York, Harcourt, Brace, 1935.
9. Lawrence, D. H., Acquired distinctiveness of cues: I Transfer between discriminations on the basis of familiarity with the stimulus, *J. exp. Psychol.*, 1945, *39*, 770–784.
10. Lawrence, D. H., Acquired distinctiveness of cues: II Selective association in a constant stimulus situation, *J. exp. Psychol.*, 1950, *40*, 175–188.
11. Spence, K. W., The nature of discrimination learning in animals, *Psychol. Rev.*, 1936, *43*, 427–449.
12. Underwood, B. J., and Hughes, R. H., Gradients of generalized verbal responses, *Amer. J. Psychol.*, 1950, *63*, 442–430.
13. Wickens, D. D., Stimulus identity as related to response specificity and response generalization, *J. exp. Psychol.*, 1948, *38*, 389–394.

Motivational Forces Underlying Learning

HARRY F. HARLOW

During the last few decades there has been a steady growth of interest in the formulation of theories designed to provide integration within and among limited psychological areas, including learning, motivation, and personality. The nature and forms of the psychological theories that have appeared during this time have been extremely diverse, and even the specific theories advanced by a single man or school have often changed radically within one or two decades. Although some psychologists may be disturbed by the evanescent trends of behavioral science theories, such trends are to be expected in a science as relatively young and unstructured as psychology today.

Even when psychologists have attempted to formulate theories of admittedly limited scope—theories primarily limited to a single area as learning, and to a single animal, as the rat—they have not yet succeeded in integrating the materials into a single, orderly arrangement which even they regard as entirely satisfactory or complete. These limitations which psychologists face in theory construction are magnified when they attempt broader theories designed to relate diverse psychological areas and the behaviors of diverse species of animals.

Whatever limitations psychological theorists face and these are admittedly many, there are advantages to be gained from attempts to interrelate such broad areas as learning, motivation, personality, and clinical research. One of the most important of the advantages lies in the fact that theories, even tenuous theories, often enable us to regard collected data from new points of view. Furthermore, new theoretical formulations invariably suggest additional experimental approaches and lead to new and different values being placed upon both past and proposed experimental programs. This organizing and evaluative function that theories inevitably have may in the long run harm or help any science. At its worst, theory may orient psychologists toward some line of research that fails to uncover new, significant factual materials or to lead to the discovery of new principles which effectively organize or interrelate any broad body of information. At its best, theory does lead to the discovery of new organizing principles and new, significant facts and functional relationships.

If we attempt to integrate learning theory and motivational theory and orient both these broad areas toward the analysis of personality, we find ourselves of necessity forced to evaluate the adequacy and importance of the earlier psychological research in the fields of learning and motivation to decide what lines of research diligently pursued in the past offer little hope of advancing personality theory, and to select neglected areas of research that offer such hope of making new contributions that they merit intensive exploration and exploitation.

We are particularly interested in the interrelated motivational and learning mechanisms having a direct role in personality formation. These complex problems lie within an area in which much of the past research has been directed toward relatively unproductive goals and in which potentially rich research possibilities have been almost totally neglected.

If we are ever to formulate an effective theory of human personality and to interrelate within this area the roles of learning and motivation, we must give proper attention and balance to all motivational mechanisms, and we must certainly not ignore the motivational mechanisms which are probably most intimately associated with human learning and the formation of human personality structure.

Motivational theory and research have in the past put undue emphasis upon the role of the homeostatic drives—hunger, thirst, sex, and elimination—as forces energizing and directing human behavior. There are, indeed, some psychological theorists who would have us believe that all or most of our adult human motives are either directly dependent upon these homeostatic drives, or are second- or third-order derived drives conditioned upon visceral needs. The fact that derived drives based on homeostatic needs are unstable and transient, the fact that the conditioned drive stimulus does not apparently reinstate the unlearned drive state [*7*, *11*], and the fact that human beings learn and live for days, weeks, or months without or in spite of a particular homeostatic need state do not disturb such psychological theorists in the least. It is, of course, the privilege of theorists to look at man from the point of view of the pylorus or to look at the pylorus from the point of view of man.

Recently, some psychological theorists, including Mowrer [*9*] and Brown [*1*], have emphasized another area of motivation—pain and conditioned pain or anxiety—and some psychological theorists have argued that these motivational mechanisms could directly or indirectly underlie much human learning and could be extremely powerful forces in shaping human personality structure. No psychologist will underestimate the importance of conditioned pain (or anxiety, if it is to be so defined), because this derived drive appears to be far less susceptible to experimental extinction than are the derived drives based on the visceral need states. But lest the role of conditioned pain and fear be overestimated, it should be pointed out that common sense tells us that the greater part of our energies are motivated by positive goals, not escape from fear and threat. Furthermore, we would emphasize that intense emotional states are theoretically unsatisfactory motives for many learned activities, particularly learned activities of a moderate or high degree of complexity. It has been recognized for subhuman animals from the time of the formulation of the Yerkes-Dodson law [*12*] that an inverse relationship exists between the intensity of motive and the complexity of task that can be efficiently learned.

It is certainly not our desire to underestimate the importance of either visceral drive states or emotional conditions as motivating forces underlying learning and influencing personality formation. Visceral drive states are important motivating mechanisms in children, and they become important motivational mechanisms in human adults under deprivation. Furthermore, the appetitive mechanisms, innate and acquired, which are associated with (even though not of necessity derived from) the visceral need states are important and persistent human motivational mechanisms. Emotional conditions including pain, fear, anger, and frustration are also important and persistent human motivational mechanisms.

But above and beyond the visceral need-appetitive motivational mechanisms and the emotional motivational mechanisms there is, we believe, a third major category of motives, a category of motives which are elicited by external stimuli and which have been described by such names as manipulation, exploration, curiosity, and play.

Psychology has doubtless suffered from the fact that these motivational mechanisms came to be labeled as "instinctive." When psychologists outlawed the term "instinct," they ruled out externally elicited drives as psychological motivating mechanisms and abandoned research in this vital area. Prior to 1915 cultural traditions had caused psychologists to repress sex and talk about curiosity; after 1915 it became popular among psychologists to talk about sex and repress curiosity.

The denial of the existence or importance of the externally elicited motives is amazing because at the common-sense level of humor and aphorism there are many references to the operation of the external-drive mechanisms. It is recognized that all primates, including man, spend a large amount of time just "monkeying around" and that monkeying around is an activity often leading to invention and creativity. There are countless cartoons bearing on the theme that the monkeys in the cage stare at the people outside and are just as amused by what they see as are the people. Köhler reported staring through a peephole to see what a chimp was doing and found that it was staring at him! Visual exploration drives in subhuman primates are clearly recognized by the saying "Monkey see, monkey do." Yet, in spite of the obvious existence of the external drives psychologists have persisted in limiting themselves to endlessly repeating, with insignificant variations, experiments designed to show the allegedly overwhelming importance of the homeostatic, internal drives. But give psychologists their due; they have at least had enough curiosity to "ape" each other's work.

Every comparative psychologist who has adapted rats for maze experiments knows that the rodents frequently run down the straightaway path used in the adaptation procedure, ignore the food, and continue to explore the environment. Frequently, the rat will refuse to eat until exploration and curiosity are sated, although it may have been deprived of food for 23 hours previously.

Furthermore, such behavior is by no means limited to rats; indeed, the prepotence of curiosity over hunger probably occurs more frequently in man than in any other animal. Those of you who have children know that if you deprive them of food for 3 or 4 or 14 hours and then seat them at the table, they will frequently engage in such activities as dropping their green beans into their milk, pouring their milk into a glass of orange juice, dangling their fork in their cup, throwing their spoon on the floor, and using their potatoes as a medium for finger painting. But in spite of the fact that the problem of dawdling at meals concerns the child psychologist, no child psychologist has ever conducted an experimental study of curiosity, manipulation, or exploration in the child. How psychologists expect to understand human motivation and the relation between human motivation and learning and personality, as long as they refuse to study one of the most basic and important motivational mechanisms that the human being possesses, must and will remain a mystery—as will also the motivational mechanism.

Fortunately, psychologists during the last few years have shown an ever-increasing interest in the experimental exploration and exploitation of the area of the externally elicited drives. Schoenfeld, Antonitis, and Bersh [*10*] have demonstrated that rats placed in a Skinner box will repetitively depress the bar even though this act produces no tangible reward. Mote and Finger [*8*] reported some years ago that rats actually decreased their running time on a straightaway maze from trial 1 to trial 2 even though no food rewards were given, and they stated that "the rats were impelled by some exploratory drive to make the running response. . . ." Keller [*6*], Zeaman and House [*13*], and Flynn and Jerome [*3*] have independently demonstrated that rats can learn when motivated only by a difference in illumination in the external environment.

During the last few years we have carried out a series of studies on the manipulation or exploration drives in monkeys and have demonstrated that monkeys learn to solve simple and complex mechanical puzzles when given no reward other than the opportunity of disassembling the apparatus. A complex six-device puzzle is shown in Fig. 1, and a learning curve for two rhesus monkeys is given in Fig. 2.

The performances of two groups of four monkeys on a five-device puzzle situation were compared by Gately [*4*]. One group was food-rewarded for puzzle solution, the other group given no extrinsic reward. The food-rewarded group learned more rapidly than the non-food group, but the best performer in the non-food group was the equal of any animal in the food group. Performance by the best member of the unrewarded group is illustrated in Figs. 3 and 4.

Recently we measured the performance of a chimpanzee on a three-device puzzle (Fig. 5). The animal was tested 20 minutes a day for 10 days. Five seconds after a puzzle was solved it was reset. Correct responses, errors, and number of puzzle completions were measured. The first 5 days' testing was conducted in early afternoon, and the last 5 days' testing between 8 and 9 P.M. Total puzzle completions and errorless puzzle completions are contained in Fig. 6, which shows that the chimpanzee performed at a high level of efficiency from day 6 on, a phenomenon also illustrated by the learning curve of Fig. 7. The puzzle problem was probably mastered on day 2, for the animal made 4 errorless solutions during the last 8 of

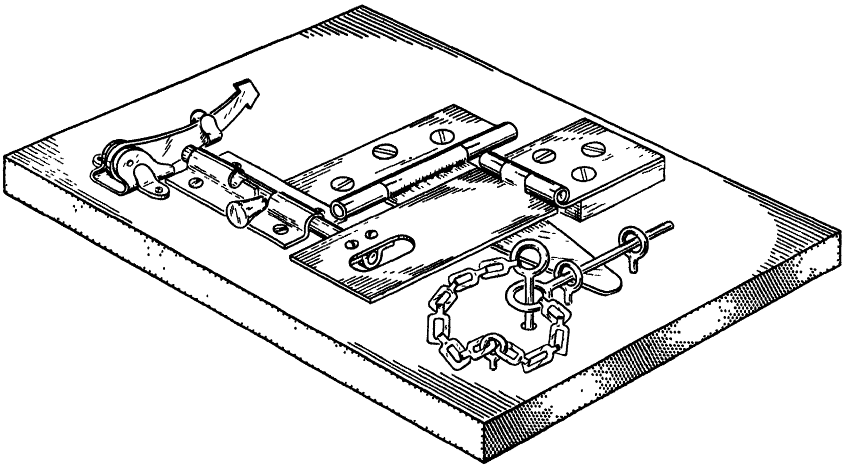

Fig. 1. Six-device puzzle.

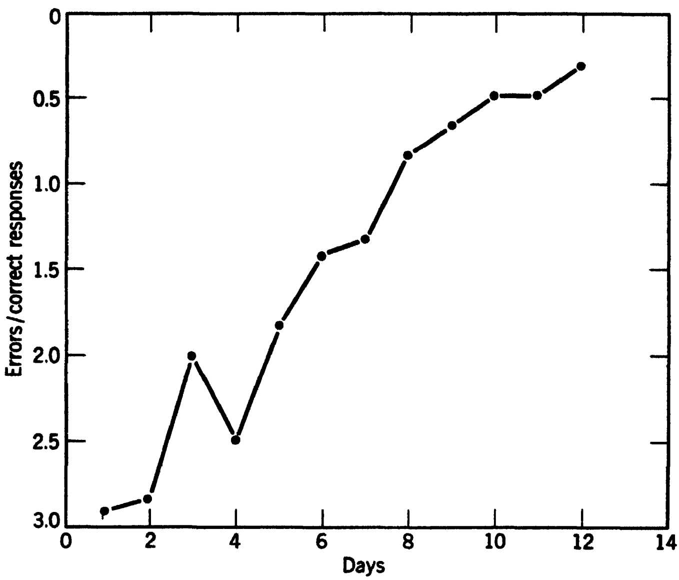

Fig. 2. Learning of six-device complex puzzle.

the 17 puzzle completions on that day. On one day the chimpanzee showed constructive exploration. After solving the problem, the animal replaced the hasp and struggled in vain to replace the hook, which is, of course, an unsolvable problem when the pin is in place. Constructive, or anabolic, behavior, we believe, will be found to be highly correlated with intellectual ability, more highly in all prob-

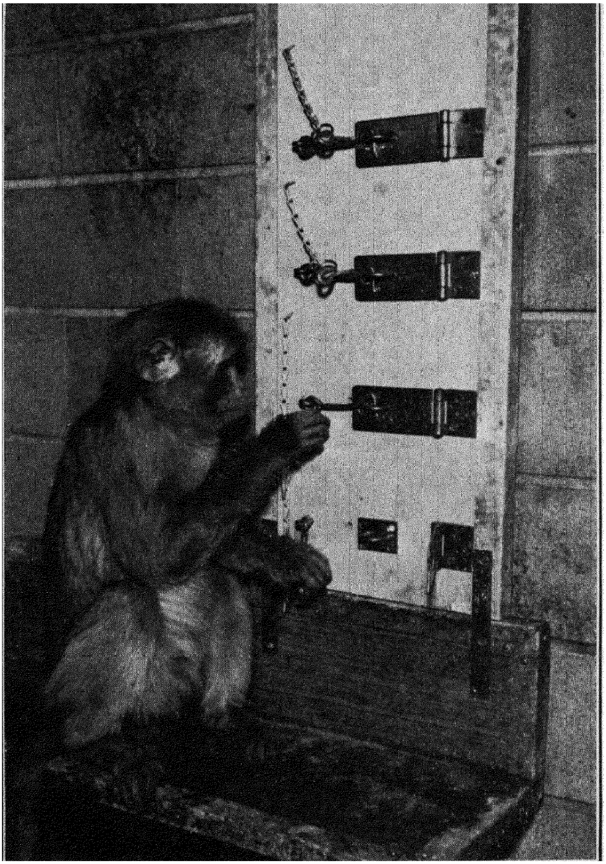

Fig. 3. Performance on multiple-device puzzle.

ability than efficiency of performance on many simple learning problems.

Harlow and McClearn [5] have demonstrated that rhesus monkeys can solve discrimination problems when no other incentives than manipulation or exploration are offered. As illustrated in Fig. 8, the monkey is presented with a panel holding five pairs of screw eyes, one member of each pair being a particular color, as red, the other a different color, as green. The screw eyes of the correct color are removable, and those of the incorrect color are fixed. Three rhesus monkeys were tested on each of seven such sets during four 5-minute trials a day for 4 days. Figure 9 shows that significant day-to-day improvement in performance on problems 1 through 5 took place, and

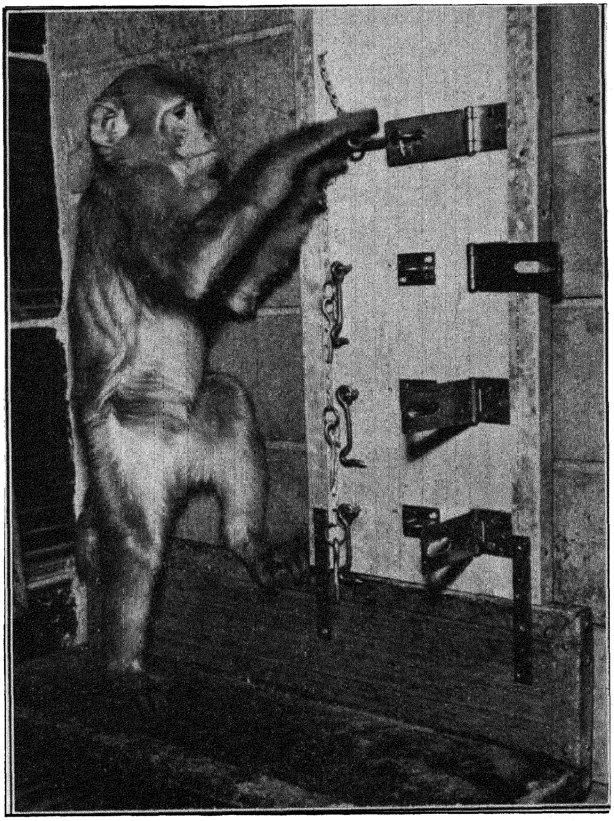

Fig. 4. Performance on multiple-device puzzle.

day 1 performance on problems 6 and 7 was superior to day 1 performance on problems 1 through 5. The mean number of responses made by the monkeys on the seven successive problems is plotted in Fig. 10. Motivation, in so far as it is measured by response frequency, is as high at the end of the experiment as at the beginning.

A series of investigations of discrimination learning by rhesus monkeys reinforced by visual exploration is being conducted at the Wisconsin Laboratory at the present time by Dr. Robert Butler [2]. The essential apparatus, a wire cage covered by an opaque box, is shown in Fig. 11. The front of the box has two hinged windows which are covered by differentially colored cards or Cellophane. The win-

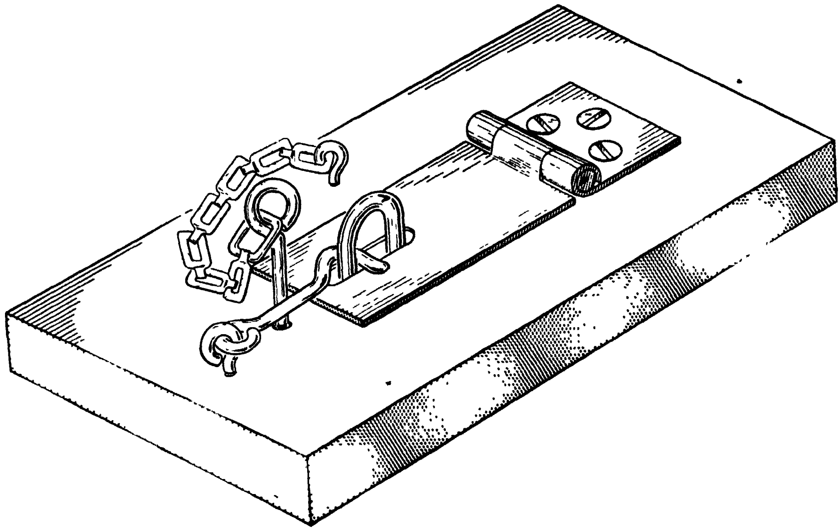

Fig. 5. Three-device puzzle.

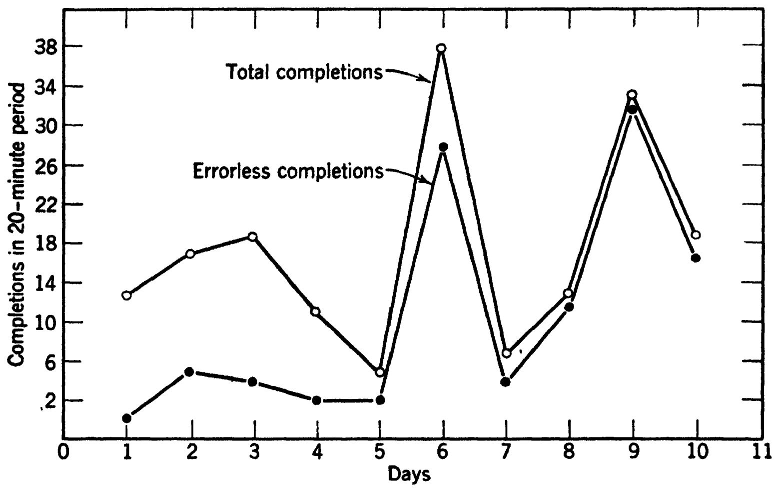

Fig. 6. Learning of puzzle by chimpanzee.

dows open outward to light pressure unless blocked by a locking device. An opaque screen which can be raised or lowered by the experimenter is immediately in front of the window inside the lightproof box. Test trial procedure is as follows. The opaque screen is raised, and the monkey is given 5 minutes to respond to the colored windows. If

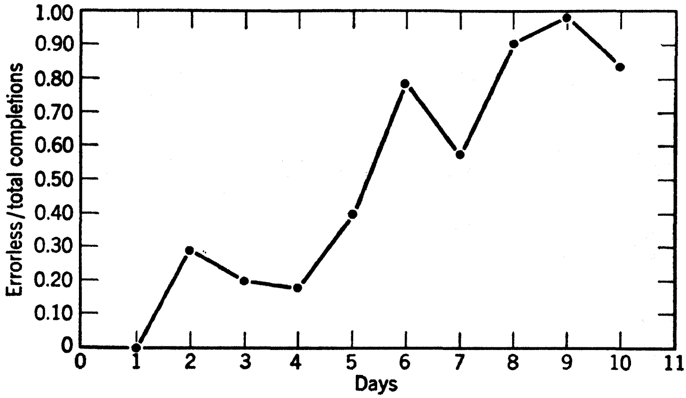

Fig. 7. Learning of puzzle by chimpanzee.

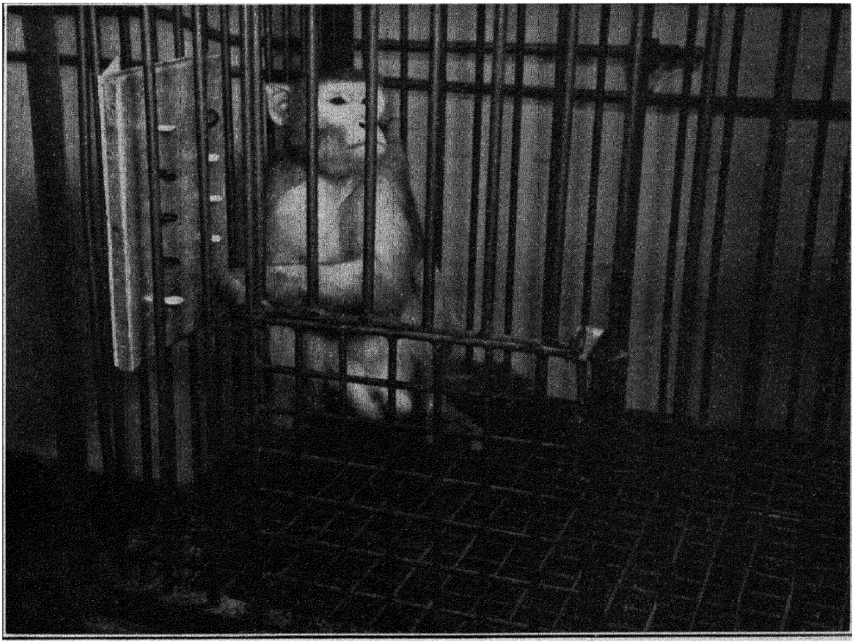

Fig. 8. Discrimination-learning problem.

an incorrect response is made, the locking device is activated and a small light outside the apparatus is turned on. The experimenter then lowers the opaque screen and waits 30 seconds before beginning the next trial. If the monkey opens the correct window, it is rewarded by being permitted to look out for 30 seconds. At the end of this

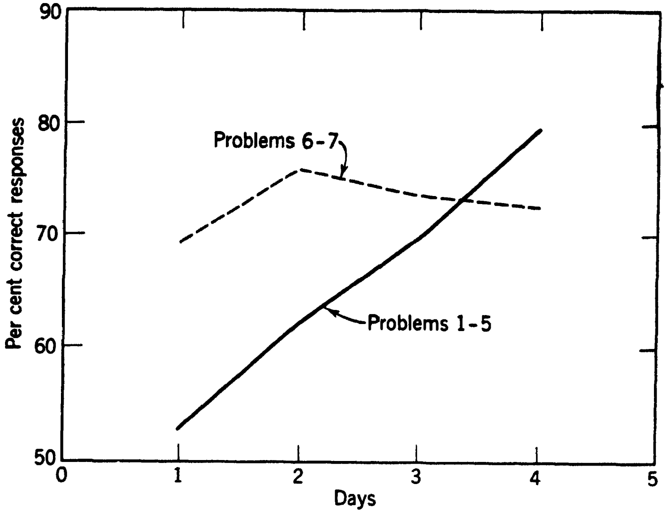

Fig. 9. Discrimination learning by monkeys.

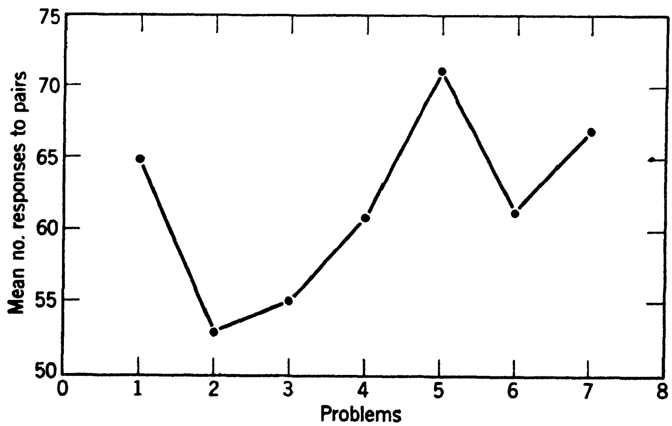

Fig. 10. Motivation as measured by response frequency.

visual reward, the opaque screen is lowered, and 30 seconds later another trial is initiated.

Discrimination learning in two rhesus monkeys tested 20 trials a day for 20 days is plotted in Fig. 12; these data show reasonably precise and efficient discrimination learning to no other incentive than visual and possibly auditory search.

The demonstration that learning to visual exploration is possible led to the investigation of the strength of this motive in rhesus monkeys. Two rhesus monkeys were tested by Dr. Butler 4 hours a day, 5 days a week, utilizing the test procedures previously described. The measure of motivational response was log latency of response from the time of raising the inner opaque screen to contact with a

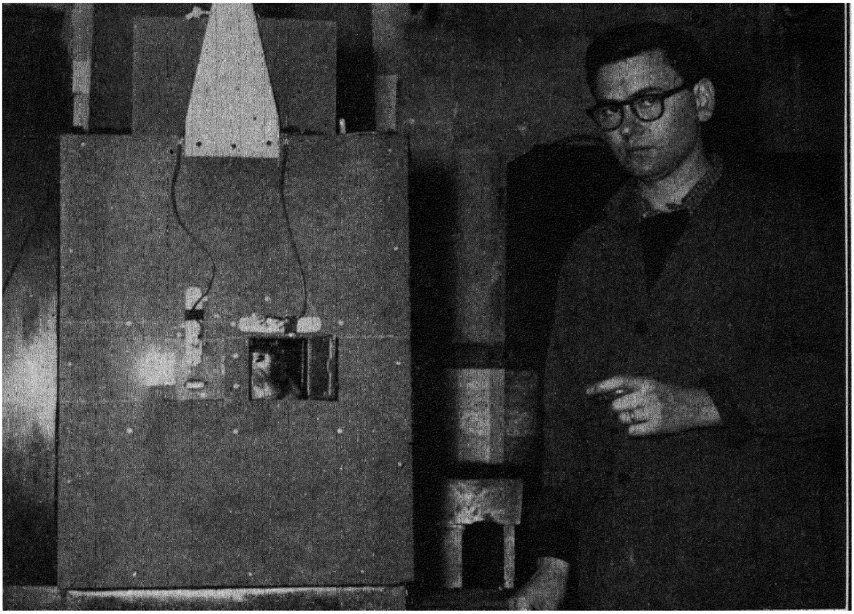

Fig. 11. Visual-exploration apparatus.

window, and Fig. 13 shows that the mean log latencies of the 5 test days did not increase during the day for either monkey if the total daily responses are divided into fourths. Performance on successive days' testing is depicted in Fig. 14. Motivational strength, in so far as it is measured by latency of response, showed an increase for one subject and no decrease for the other.

Although learning was investigated only incidentally in this study, the data presented in Fig. 15 demonstrate rather rapid learning and extremely proficient performance subsequently in monkey 102. The other subject, 147, apparently learned, as indicated by long series of errorless runs, but did not maintain highly consistent performance over any block of 100 trials.

Satiation tests have been run by Dr. Butler on two monkeys thus far, "satiation" being defined as failure by the monkey to respond to

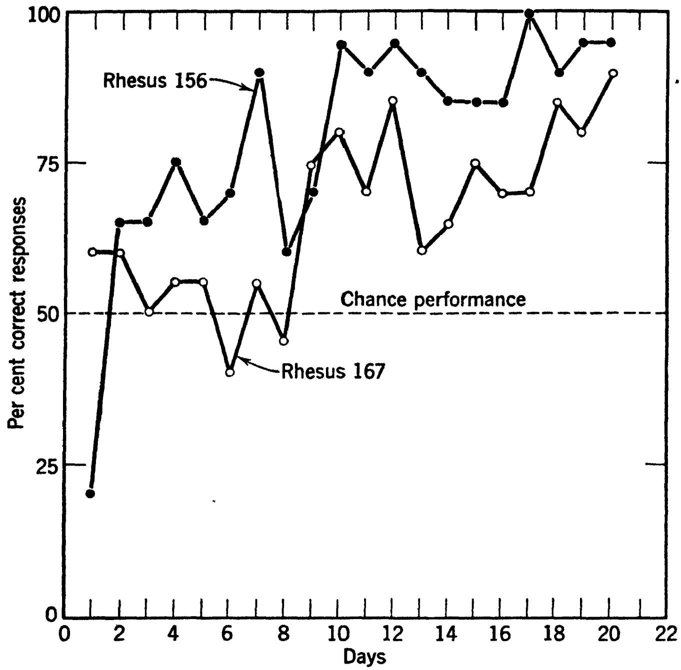

Fig. 12. Discrimination learning to visual exploration.

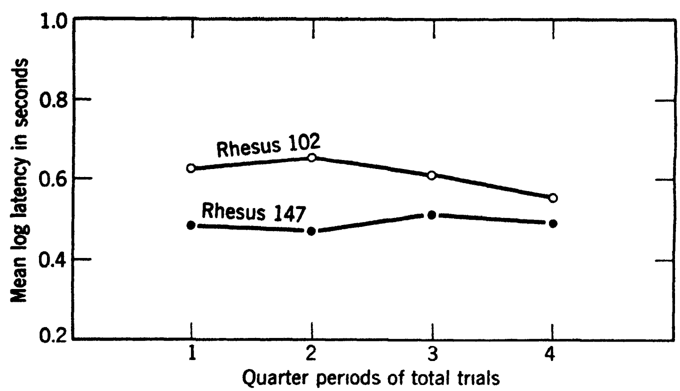

Fig. 13. Visual motivation as measured by response latency.

a window within the allotted 5-minute period on two successive trials. The persistence of the visual exploration motivation far exceeded predictions. One monkey responded continuously for 9 hours, and the other for 19.5 hours, as shown in Fig. 16.

Initial steps toward the quantification of the factors underlying visual exploration have recently been undertaken by Dr. Butler. The

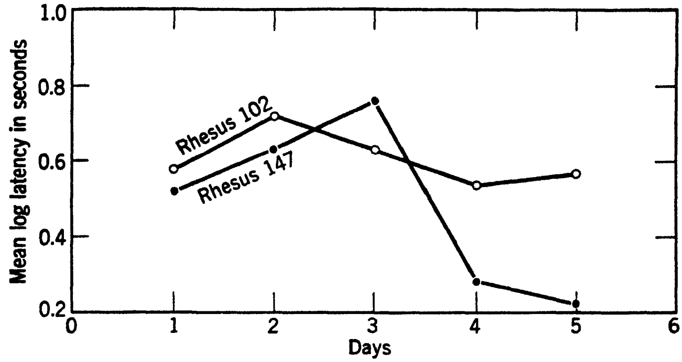

Fig. 14. Response latency as a function of days.

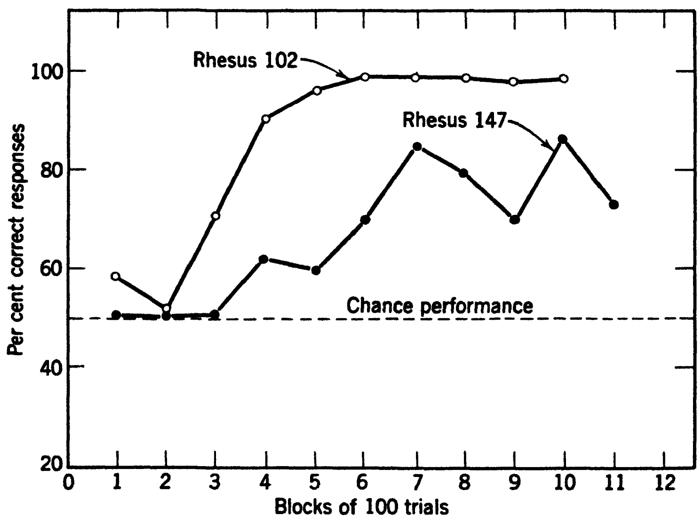

Fig. 15. Learning to visual exploration.

basic apparatus used in these researches differs from that illustrated in Fig. 11 only in the fact that there is a single, hinged door, centered in the face of the apparatus, rather than two doors. Facing and connecting the front of the basic apparatus is a box 48 × 36 × 30 inches with top and sides covered by heavy black cloth and illuminated by a 100-watt lamp.

Each trial was initiated by raising an opaque screen which exposed the single window; when the monkeys pushed open the window, they were permitted to explore the illuminated box visually for 5 seconds. Four observational conditions were measured: (1) an empty box control condition, (2) an array of five foods, (3) a moving toy electric train consisting of an engine and two cars, and (4) another monkey confined in a transport cage.

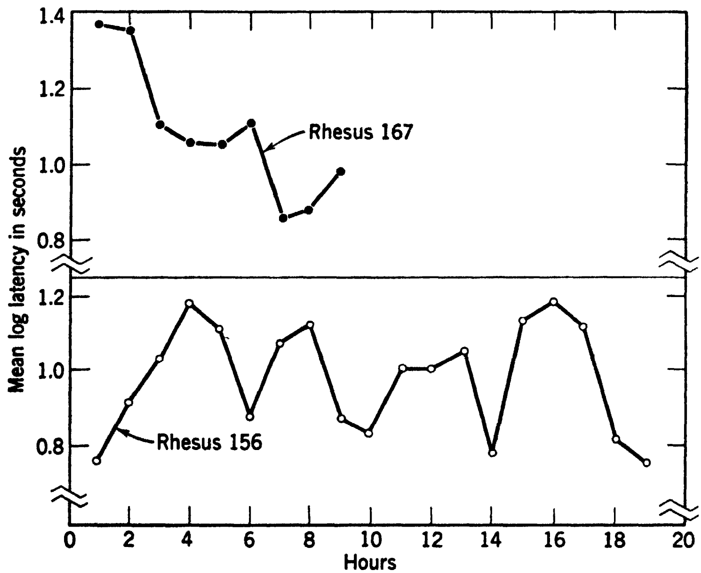

Fig. 16. Satiation of visual-exploration drive.

Eight rhesus monkeys were tested 30 minutes a day for 20 days. Each incentive condition was continued for 5 days, and order and frequency of incentive conditions were balanced.

As indicated in Fig. 17, strength of incentive as measured by frequency of responses ranged downward from monkey, to train, to food, to empty control. Frequency of response to the monkey and train observational conditions was significantly greater than frequency of response to the empty control condition at the 0.01 per cent confidence level. Motivational strength, as measured by the total daily number of responses, showed no decreases throughout the experiment.

Now I am certain that every learning theoretician present has come to realize that the facts concerning exteroceptive motivation are in complete accord, or almost complete accord, with his theoretical position. I am certain that Dr. Spence can describe how they fit drive-reduction theory, and I have tried to make this easy by speaking of manipulation drive and visual exploration drive, in spite of reservations I have about the appropriateness of the term "drive" in this context. I am equally certain that Dr. Tolman could explain how our data fit within the framework of expectancy theory, and I am certainly certain that Dr. Guthrie could explain how they fit Guthrie

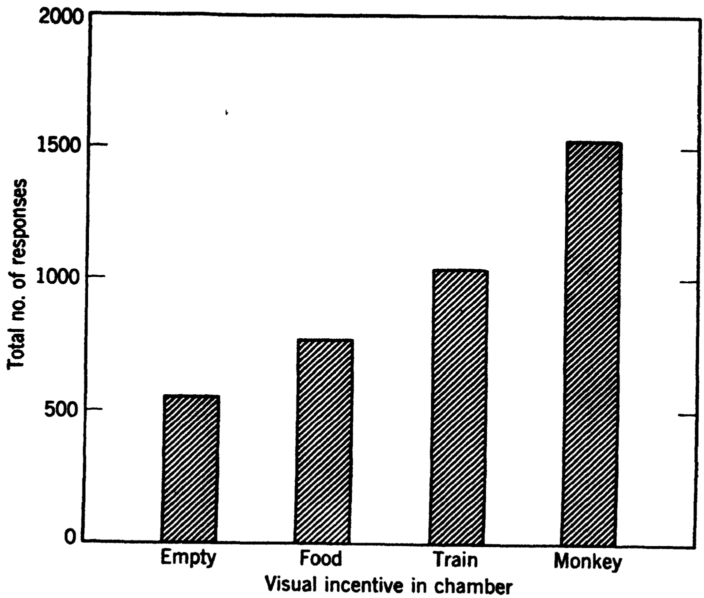

Fig. 17. Factors influencing strength of visual-exploration drive.

theory—because they really do fit the basic motivational principle of Guthrie theory. However, I do not believe, and do not intend to believe, Guthrie theory no matter how well the Wisconsin data support the position. More explicitly, I endorse the Guthrie position that animals learn responses regardless of the locus of the stimulating agent, but beyond this my own theoretical position has little in common with that of Guthrie. Perhaps our data will illustrate a psychological law, and this psychological law is one that has great generality. Psychological theories seldom unify facts, but facts often unify psychological theories.

Before the professional theorists explain our data to us, however, I should like to point out a number of facts. It appears to be most unlikely that exteroceptive motives can be explained as second-order, derived motives conditioned to hunger, thirst, or any other primary drive. The manipulatory and visual exploratory drives are extremely persistent, whereas derived drives conditioned to so-called hunger or thirst show relatively rapid extinction. Indeed, one of the important theoretical aspects of the manipulatory and visual exploratory drives resides in the fact that they may show increasing motivational strength rather than decreasing motivational strength with repeated elicitation. This phenomenon would make it necessary for us to postulate some such phenomenon as Inverse Experimental Extinction if we are to account for manipulatory and visual exploration drives in terms of derived motives.

We are convinced that the externally elicited motivational systems are as fundamental and as innate as are the hunger-appetite and thirst-appetite systems. By this we do not mean that learning does not operate as a component in the externally elicited motivational systems of the adult animal; we merely mean that the learning component here is probably no larger than it is for the hunger-appetite motivational system. We firmly believe that externally elicited motivational systems interact with the hunger-appetite and thirst-appetite systems, but we do not believe, and there is no evidence to support the position, that any of these systems is derived from the other, or that any differential degree of dependence exists among them.

Recently, motivation theorists and personality theorists have shown some motivational obsession about anxiety. No one will deny that anxiety, or many other emotional states, serve as motives—but any assumption that anxiety has some special, prepotent, motivational role has yet to be established. At best, anxiety is a motive for avoidant behavior, and the greatest part of human motivation is positive searching toward goals, not mere avoidance. In spite of our faith in the importance of positive, forward-oriented motives such as curiosity, manipulation, and exploration, we do not wish to put any constraints, in research or theory, upon the anxious psychologists. We merely wish to insist, however, that even if some psychologists are scourged to their experimental dungeons like quarry slaves, the remainder of the population will continue to be motivated by pleasant and positive incentives.

REFERENCES

1. Brown, C. J., Learnable drives, Nebraska Symposium, 1952.
2. Butler, R. A., Discrimination learning by rhesus monkeys to visual-exploration motivation, *J. comp. physiol. Psychol.*, 1953, *46*, 95–98.

3. Flynn, J. P., and Jerome, E. A., Learning in an automatic multiple-choice box with light as incentive, *J. comp. physiol. Psychol.*, 1952, *45*, 336–340.
4. Gately, M. J., Manipulation drive in experimentally naïve rhesus monkeys, M.A. thesis, University of Wisconsin, 1950.
5. Harlow, H. F., and McClearn, G. E., Object discriminations learned by monkeys on the basis of manipulation motives, *J. comp. physiol. Psychol.*, 1953, in press.
6. Keller, F. S., Light aversion in the white rat, *Psychol. Rec.*, 1941, *4*, 235–250.
7. Miles, R. C., and Wickens, D. D., Effect of a secondary reinforcer on the primary hunger drive, *J. comp. physiol. Psychol.*, 1953, *46*, 77–79.
8. Mote, F. A., and Finger, F. W., Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response, *J. exp. Psychol.*, 1942, *31*, 57–69.
9. Mowrer, O. H., *Learning theory and personality dynamics*, New York, Ronald Press, 1950.
10. Schoenfeld, W. M., Antonitis, J. J., and Bersh, P. J., Unconditioned response rate of the white rat in a bar-pressing apparatus, *J. comp. physiol. Psychol.*, 1950, *43*, 41–48.
11. Simon, C. W., Wickens, D. D, Brown, U., and Pennock, L., Effect of the secondary reinforcing agents on the primary thirst drive, *J. comp. physiol. Psychol.*, 1951, *44*, 67–70.
12. Yerkes, R. M., and Dodson, J. D., The relation of strength of stimulus to rapidity of habit formation, *J. comp. Neur. Psychol.*, 1908, *18*, 459.
13. Zeaman, D., and House, B. J., Response latency at zero drive after varying numbers of reinforcements, *J. exp. Psychol.*, 1950, *40*, 570–583.

The Premature Crystallization of Learning Theory

NORMAN R. F. MAIER

In recent years a great deal of emphasis has been placed on developing a learning theory, and we may actually classify certain psychologists as learning theorists. Many of our courses in psychology now devote more time to discussing the relative merits of various learning theories than to the facts of learning, and our graduate students are required to develop skills in applying several theories to a given set of data. Thus the student's knowledge often is judged by how well he knows what different psychologists think and not by how well he is acquainted with experimental subject matter. Granting that both theories and experimental facts are important, what constitutes a healthy balance? I personally feel that an interest in theories is desirable for the development of science because theories help us organize facts and they help us to ask good research questions. However, an interest in theories can become a liability if it prevents us from exploring certain kinds of relationships or causes us to ignore facts that do not fit the theory with which we identify ourselves. When these things occur, the theory becomes an attitude and ideas become good or bad rather than right or wrong.

Perhaps we are somewhat overambitious and have assumed that psychology is more advanced than facts warrant. We seem to want a learning theory that works not only for all learning situations but also for all behavior. We seem to want to predict, to do research by stating hypotheses, and seem no longer to be content with asking questions of the universe and getting our answers through research. When we ask questions we can be open-minded and are likely to let the answer to one question influence the nature of the next. It is in this way that we get acquainted with our universe. However, when we predict we show our maturity, and we can even determine a scientist's success by calculating his percentage of correct predictions. But we must not act more mature than we really are.

We want a learning theory to live by, despite the fact that maze learning, conditioning, discrimination learning, and mastering a problem box are uncorrelated functions and despite the fact that various learning theorists developed their viewpoints by studying the learning of particular animals in situations which were rather specific and which delimited the animals' responses.

It would seem wiser if we would generalize somewhat less and see whether agreement could be reached if we did not mix data obtained from such a great variety of situations. Thus the fact that an animal shows insight in a maze does not mean it must behave insightfully in a puzzle box. May not random behavior and insightful behavior be different in kind? At least we should not make an assumption one way or the other until such a question has been adequately explored and evaluated.

In a previous paper [6] I tried to show that there are at least five unrelated functions which operate in learning situations and that these functions are involved to different degrees in various learning tests. These functions are variability, plasticity, perception, behavior repertoire, and association fixation. (These functions are in addition to reasoning [7] and frustration-induced fixations [*11*, *12*, *13*] which I exclude from learning proper.) Since a test score measures the resultant of the various possible functions it would seem unwise to formulate theories which assumed the score to be a measure of a single function. Such an assumption discourages exploration of basic issues about what goes on in the animal when he solves a problem.

To illustrate how a well-developed theory may lead to oversimplification let us take a brief look at the treatment of reasoning given by Dollard and Miller [3, p. 111]. The problem has to do with a person caught in a traffic jam who needs to make a left turn onto a crowded highway. He notices that cars coming toward him have no trouble making a right turn onto the highway. The driver is caused to say to himself, "If I were only going the other way, it would be so easy!" Once this statement is made the problem is one which the driver has had "a great deal of practice in solving." He merely gets into the right-hand lane, drives past the traffic jam, turns around at an uncrowded intersection, comes back the other way, makes a right turn, and reaches his objective.

I feel that two things have been ignored by this analysis. First, what causes a person to ask himself *good* questions? Many other responses can be elicited by a situation of this sort. Surely a selective process must operate, and a selection of the right question by random trial and error would be a long process. As a matter of fact, almost all the research on reasoning has hinged on the study of this *selection* mechanism which Dollard and Miller by-passed. Second, does the solution to the problem used require the highest mental process or might reproductive thinking be adequate? Certainly it is not the kind of problem used by investigators who consider the reasoning process more complex than the learning process.

Dollard and Miller's treatment of fixated behavior reveals a similar circumvention. In a footnote devoted to this subject they point out that confusion arises in interpreting the fixated behavior of my rats because of failure "to note that they were rewarded on every trial (irrespective of whether they jump to the correct window or not) by escaping from the punishing air blast." This statement assumes that I could not explain my rat's persistent incorrect behavior in a discrimination problem and formulated a non-reinforcement theory, without recognizing that escape from an air blast was a reward. As a matter of fact this point was too obvious for me to overlook. I even speculated over some complex probability notions suggested by Humphreys [*4*, *5*]. However, the phenomenon of persisting behavior is not the fact that gave rise to my theory. The case for fixated behavior, right from the beginning [*13*], was based upon (*a*) the appearance of a different number of rats with fixated responses in the two training conditions used (both of which involved the escape from the air blast mentioned above); and (*b*) the appearance of bimodal distributions in each of the groups. Since then additional facts have been reported, but these followed the initiation of the theory.

CURRENT RESEARCH FINDINGS INDICATING NEED FOR BROADER VIEWPOINT

I should like to take this opportunity to bring together some studies which I believe tend further to point up the need for exploring wider areas before jelling too specifically on a learning theory. I believe that these experiments bear on learning theory but would not be suggested by learning theory. In this sense I am making a case for keeping a learning theory more open-ended. One must leave room for empiricism and be willing to work with what Conant [*1*] calls "fuzzy" ideas. He points out that "clear-cut operational definitions are never possible in the infancy of a science" [p. 73] and suggests that they may be a handicap.

The experimental work on abnormal fixation has suggested a kind of behavior which does not follow learning principles as they are now conceived. This observation was in part based upon varied evidence which when combined indicated that position responses developed under frustrating conditions are more difficult to change than position responses developed under conditions of motivation or reward training. Recently we retested this conclusion, using a large number of rats which matched responses but developed under the two different conditions.

A total of 49 rats were rewarded with food for choosing the right- (or left-) hand card and punished for choosing the left- (or right-) hand card in a discrimination apparatus. After nearly 160 reinforcements the position response was firmly established and the rats were required to learn a card response. It was found that after 200 trials 34.7 per cent learned the card discrimination in from 30 to 150 trials, 49.0 per cent persisted in the position response, and 16.3 per cent showed variable behavior.

A total of 55 rats developed position responses under insoluble problem conditions. After a similar number of trials they too were required to learn a card response. Only 5.5 per cent were able to learn the card discrimination, 92.7 per cent persisted in the position response, and 1.8 per cent showed variable behavior. That this difference was not due to some difference in the learning or need condition of the animals in the two groups was shown by the fact that animals failing to adopt the card-discrimination response nevertheless learned to discriminate between the two cards. They expressed this learning by jumping directly to the positive card and abortively to the negative card. Furthermore, the differential abortive jumping was shown to an equal degree by animals with and without position fixations. Thus fixated behavior does not prevent learning from taking place but rather prevents certain learned behavior from becoming expressed.

We previously concluded, on the basis of this and other evidence, that fixated behavior is of a compulsive nature. If one assumes that frustration-induced fixations have a compulsive property it follows that learning which does not violate the compulsive tendency will be expressed more readily than learning which violates it.

Mr. Paul Ellen is exploring this possibility. After an animal has developed a persistent position response and has failed to adopt a card-discrimination response, but has expressed its discrimination by jumping abortively to the negative card and directly to the positive card, it is confronted with a three-window situation. Suppose that a particular animal always jumps to the right window of a pair when the jumping stand is directly in front of the two windows. (This jump to the right occurs even when the positive card is on the left and the negative card is on the right.) If we now confront such an animal with three windows and place the jumping stand directly in front of the first two, then there will be two windows to the animal's right. If the first, second, and third windows contain positive, negative, and positive cards, respectively, it is possible for the rat to make a jump to the right and to the positive card at the same time (by jumping to the extreme right). By using various combinations of three windows in this manner Mr. Ellen was able to cause a fixated rat to follow the positive card as it was moved between the second and third positions. Eventually the fixation was broken when the rat was able to follow the positive card to the left. Thus by training the animal to respond within the bounds of its fixation it became possible to break some fixations at least.

It seems that the whole problem of persistent behavior must be carefully explored with a more empirical approach. There are perhaps many kinds of rigidity,* and attempts to account for compulsive behavior in terms of our present learning concepts may make them unacceptable to clinicians.

In recent years I have made much of the point that frustrating or insoluble problem situations divide a group of individuals [*11*, *12*],

* Possible mechanisms giving rise to persisting or rigid behavior are the following:

(*a*) Relatively greater strength of response in question over other responses in organism's repertoire.
(*b*) No motivation to change is present in situation.
(*c*) No other responses in organism's repertoire.
(*d*) Motivation to persist is present in the form of fear of change.
(*e*) Compulsiveness created by frustration.

The postulation of a variety of persistent behaviors also assumes basic differences in procedures for correcting them. These corrective procedures may take the following forms:

(*a*) Increasing the habit strength of alternatives by repeated reinforcement.
(*b*) Altering the animal's needs (change from hunger to thirst), which should cause shift in behavior even if situation is unchanged.
(*c*) Exposing animal to additional learning or knowledge (e.g., exploration of new territory).
(*d*) Removal of factors giving rise to fear (e.g., a permissive environment).
(*e*) Removing state of frustration, guidance, and other procedures under investigation.

be they rats or college students, into two populations: (*a*) those whose subsequent learning scores fall within the range of a normal distribution curve, and (*b*) those whose scores fall outside the distribution curve and show them to be clearly handicapped in their learning. We explain this break in a group by saying that a given situation frustrates some individuals and not others. Thus when an individual's frustration threshold is exceeded he moves from one distribution curve to another. Marquart [*14*] previously reported that a period of random punishment (electric shocks) on 75 per cent of the trials in an insoluble problem disturbed the subsequent learning scores of college students when the problem was made soluble. Seventy-four per cent of the students performed like a control or an unpunished group, but the other 26 per cent formed a new population and made an average score more than 4 times that of the control group.

Recently Marquart and Arnold [*15*] repeated this experiment with another group of students, but instead of electric shock punishment they used a light signal to indicate an incorrect choice. Again a bimodal distribution of scores was obtained when learning on a soluble problem was tested. The number of individuals falling near one mode or the other was related to the extent of failure experienced by three different groups, but the point at which there was a gap in score between the two modes was the same. In this instance the proportion of normal to extreme performers was 77 to 23 per cent, or almost exactly the same as in the first study. Thus a period of failure, not punishment as such, causes about one-fourth of a group of persons to behave qualitatively differently from the other three-quarters.

That the presence or absence of frustration is a determining factor in whether or not a given individual performs normally or exceedingly poorly was suggested both by the responses of the individuals during the experiment and by the relationship of extreme scores with certain personality variables. Certainly an attempt to explain the results in terms of differences in learning aptitude is unrealistic.

In another experiment with animals, Ellen and I tested the relative influence on learning of variations in reinforcement and of variations in the opportunity to perceive a change in effect. Animals were trained to form a position response. After 160 trials an attempt was made to replace this response with a card-discrimination response. Three training methods were selected which seemed adequate to evaluate the reinforcement and perceptual functions.

Method A. The positive card placed on the side of the previously learned position response on 8 out of the 10 trials given daily and the negative card on the position side for the other 2 trials. This method results in a *reward:punishment* ratio of 80:20 as long as the animal continues to express a position response.

Method B. The positive card placed on the side of the position response in 5 out of each 10 trials and the negative card on the position side for the other 5 trials. This method results in a *reward:punishment* ratio of 50:50.

Method C. The positive card placed on the side of the position response in 2 out of each 10 trials and the negative card on the position side on 8 trials. This method results in a *reward:punishment* ratio of 20:80.

These three methods differ in the degree of reinforcement given the position response as well as in the absolute number of times that a card response is reinforced. Thus, regardless of how one wishes to treat reinforcement, including partial reinforcement [*18*], one should expect animals to be able to learn the new response as well as give up the old one. Further, Methods *A* and *C* should be more different from each other than either is from Method *B*.

If, however, the perception of inconsistency is important to learning, then Methods *A* and *C* should be more like each other than either is like Method *B*. By Methods *A* and *C* the same thing happens on nearly all trials, either almost always reward for a position response or almost always punishment for a position response. Thus the world is fairly orderly and consistent, which means that the animal's expectations are nearly always fulfilled. The most inconsistent condition is Method *B*, when either effect (reward or punishment) is just as probable as the other. Should this irregularity inspire learning, in that the need for learning is great, or might it be discouraging?

The results obtained are shown in Table 1. First, it is clear that none of the conditions assured learning, although there was more than ample time for this. Thus a given training procedure teaches some animals and entirely fails with others. Such wide differences in learning performance cannot be set aside by assuming that a bimodal distribution of aptitudes exists.

Second, since the results of Method *B* do not lie between those of Methods *A* and *C*, it is apparent that there is no trend which follows the degree of reinforcement. Rather, Methods *A* and *C* are strikingly similar with regard to both the number of animals abandoning the position response and the number adopting the card response. If the results of Methods *A* and *C* are combined they are significantly different from those of Method *B*. ($P = 1$ per cent for abandoning the position response and between 1 and 2 per cent for adopting the new response.) Thus it seems that the method yielding the most inconsistent effect caused the fewest number of animals to change.

One might suppose that this failure to cause so many animals to learn is due to the fact that Method *B* is more difficult than the other two. However, the learning data (last column) exclude this possibility. Not only are there no significant differences in learning scores,

TABLE 1 *Effect of three reward-punishment ratios on behavior modification*

	R:*P* Ratio	No.	No. Abandoning Position Response	No. Adopting Card Response	Trials to Learn
A	80:20	15	9	7	114.3
B	50:50	20	6	4	90
C	20:80	14	10	6	101.6

but also the trend suggests that condition *B* makes for faster learning. Method *B* thus seems both to facilitate learning and to disturb learning more than the other methods, and which effect it produces appears to be a function of the individual. It is for this reason that insightful problem solving, trial and error, and frustration-instigated behavior cannot be described in situational terms. Any of these behaviors can be stimulated by a situation in which there is thwarting or blocking of goal-oriented behavior [*10, 11*]. What occurs in a given individual depends on such factors as (*a*) physiological differences, which may merely be threshold differences in emotion; (*b*) personality differences, which may reflect different degrees of determination; (*c*) intellectual differences, which may determine the relative difficulty of the problem for the individual; (*d*) perceptual differences, which may be determined not only by past experiences but also by chance mental sets and minor neurological differences; (*e*) need differences, which vary from individual to individual as well as with the period of deprivation; and (*f*) the repertoire of learned responses. Because all these processes influence what an individual learns or how he changes as a consequence of an exposure to a situation they are relevant to learning, yet only the last two are generally given full recognition.

In previous studies from our laboratory [2, *6*, 8, 9, *17*] reference has been made to perception as a determiner of what is learned. Experimental data [*8*, 9, *17*] previously published showed that various rats in mastering the same discrimination problem actually came away with different learning, as tested by the method of equivalent stimuli.* The same studies also showed that the difficulty of a discrimination problem, as well as what was learned, differed, depending upon which of a pair of stimuli was made positive.

It seems easier to handle these variables and relationships in a perception theory than in a learning theory, and I feel that present learning theory discourages research which investigates these matters. However, today I wish to go even a step farther in complicating matters for learning theory and indicate how the selection or expression of behavior can be influenced by an emotional tone.

Zucker [*19*] has demonstrated that delinquent children completed some of his stories differently from non-delinquent children. William Edwards, one of our graduate students, gave Zucker's completion stories to three groups of sixth-grade children (mean age 12.6 years) in a role-playing situation. One group was instructed to feel rejected by their parents, another group was instructed to feel wanted or accepted by their parents, and the third group was not instructed.

The results obtained are presented in Table 2 along with Zucker's data. From this table one can see that the story-completion results for the "rejected" group were significantly different from those for the "accepted" group, whereas the uninstructed group made scores part way between those of the other two groups but somewhat closer to the scores of the "rejected" group.

Further, the results of the "rejected" group were strikingly similar to those of Zucker's delinquent group, and the "accepted" group completed stories similar to those of Zucker's non-delinquent group. When the data for the four stories are combined we find that, on the average, 76 per cent of the "rejected" group gave delinquent responses, which compares closely with Zucker's 75 per cent for actual delinquents. Only 32 per cent of the "accepted" group gave delinquent responses, which is slightly higher than the 24 per cent obtained from Zucker's non-delinquents.

* Differentiation between the stimulus objects could be made in terms of the absolute properties of either (or both) the positive or negative stimulus, relative over-all brightness, figure size, ground properties, etc., and the relative use made of these differences varied greatly from individual to individual.

Edwards' population was chosen from a public school in which the deliquency rate was high. This perhaps explains why the uninstructed group (control) approached the group who were instructed as "rejected" and why the "rejected" group duplicated Zucker's results with a group of actual delinquents somewhat more accurately than the "accepted" group duplicated his non-delinquent group.

It seems that this study, as well as some other role-playing results [*16*], indicates that the expression of behavior can be controlled by manipulating attitudes or feeling states. This is indicated by the dif-

TABLE 2 *Results of role playing, delinquency, and non-delinquency*

		Zucker Data per cent		Role-Play Data per cent		
	Delinquent Solution	D	ND	Rejected	Accepted	Control per cent
Number tested		25	25	35	34	26
Story I	Went away	88	28	83	29	91
II	Kept knives	68	24	83	26	64
III	Stole	76	24	60	32	50
IV	Went to friend	68	20	77	41	64
Average		75	24	76	32	67

ference in the story-completion results of the two instructed groups. More important, however, is the observation that Edwards' children seemed to have both delinquent and non-delinquent behavior in their repertoires. Ordinarily we speak of training delinquents to be good citizens and assume that children learn delinquency. Without introducing reinforcement and without giving them experience with relevant behaviors, different responses were elicited from two groups of children when only the role was changed.

Surprising too is the fact that these children "seemed" to know that delinquency and parental rejection went together. Where did they learn this? Might it not be better to say that when they felt rejected they perceived the world as hostile and unfriendly and that the behavior elicited by these perceptions is known as delinquent behavior? I personally feel that feelings of frustration produce behaviors which do not have to be learned. However, they can be learned and they can be modified by learning, but learning is not a basic essential for a good deal of the behavior which is expressed.

To assign learning a greater role in behavior than it actually plays is to exclude consideration and exploration for other variables.

CONCLUSIONS

Theory formulation is not necessarily a sign of progress, and it may actually be a disservice to science if it becomes an end in itself. The type of research reported in this paper suggests that the concept of reinforcement, so important to present learning theory, requires re-examination. Any attempt to incorporate association formation, motivation, and perception into a single quantitative theory seems premature, since our knowledge of each of these processes still is in a state of development. To combine them all into a reinforcement concept buries the problems rather than stimulates analysis.

Progress in science can be made by developing principles and doctrines to describe qualitative relationships, and formulating laws to describe quantitative relationships. Such objectives of research permit the exploration of ideas and encourage speculation and the testing of hypotheses. They permit one to think of learning in connection with the counseling process and group-discussion processes as well as in connection with training experiments. Laws and doctrines may be amended, modified, extended, or refuted with relative ease since these changes do not threaten a whole system. Not only are the fruits of such research stimulating to future research, but also in the meantime they may serve practical uses in education, character development, and therapy.

The history of science has been one of gradual clarification of ideas. Theories of limited scope make their appearance, and eventually some of them may incorporate others or give rise to a larger theory. Psychology is at the stage where theories of perception, of motivation, of association formation, of organization, of reasoning, and of frustration can be formulated so as to organize large bodies of knowledge. However, discrepancies and disagreements will be numerous.

A general behavior theory is premature at the present time because it (*a*) discourages exploratory research; (*b*) emphasizes quantitative measurement at the expense of qualitative analysis; (*c*) assumes that science develops along deductive logical lines, thereby excluding many other sources of development and kinds of thinking; and (*d*) is unwilling to entertain concepts which are still vague and in the process of development.

REFERENCES

1. Conant, J. B., *Modern science and modern man,* New York, Columbia University Press, 1952.
2. Eglash, A., Perception, association, and reasoning in animal fixations, *Psychol. Rev.,* 1951, *58,* 424–434.
3. Dollard, J., and Miller, N. E., *Personality and psychotherapy,* New York, McGraw-Hill, 1950.
4. Humphreys, L. G., The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions, *J. exp. Psychol.,* 1939, *25,* 141–158.
5. Humphreys, L. G., Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *J. exp. Psychol.,* 1939, *25,* 294–301.
6. Maier, N. R. F., The specific processes constituting the learning function, *Psychol. Rev.,* 1939, *46,* 241–252.
7. Maier, N. R. F., The behavior mechanisms concerned with problem solving, *Psychol. Rev.,* 1940, *47,* 43–58.
8. Maier, N. R. F., Qualitative differences in the learning of rats in a discriminating situation, *J. comp. Psychol.,* 1939, *27,* 289–331.
9. Maier, N. R. F., The effect of cortical injury on equivalence reactions in rats, *J. comp. Psychol.,* 1941, *32,* 165–189.
10. Maier, N. R. F., Reasoning in humans: III The mechanisms of equivalent stimuli and of reasoning, *J. exp. Psychol.,* 1945, *35,* 349–360.
11. Maier, N. R. F., *Frustration: the study of behavior without a goal,* New York, McGraw-Hill, 1949.
12. Maier, N. R. F., and Ellen, Paul, Can the anxiety-reduction theory explain abnormal fixation? *Psychol. Rev.,* 1951, *58,* 435–445.
13. Maier, N. R. F., Glaser, N. M., and Klee, J. B., Studies of abnormal behavior in the rat: III The development of behavior fixations through frustration, *J. exp. Psychol.,* 1940, *26,* 521–546.
14. Marquart, D. I., The pattern of punishment and its relation to abnormal fixation in adult human subjects, *J. gen. Psychol.,* 1948, *39,* 107–144.
15. Marquart, D. I., and Arnold, L. P., A study in the frustration of human adults, *J. gen. Psychol.,* 1952, *47,* 43–63.
16. Solem, A. R., The influence of the discussion leader's attitude on the outcome of group decision conferences, Doctoral dissertation, University of Michigan, 1952.
17. Wapner, S., The differential effects of cortical injury and retesting on equivalence reactions in the rat, *Psychol. Monogr.,* 1944, *57.*
18. Wilcoxon, H. C., "Abnormal fixation" and learning, *J. exp. Psychol.,* 1952, *44,* 324–333.
19. Zucker, H., The emotional attachment of children to their parents as related to behavior and delinquency, *J. Psychol.,* 1943, *15,* 31–40.

Learning and Explanation

DONALD K. ADAMS

A sentiment or a psychical system or a superego is just as physical a notion as a cell assembly, an engram, a reflex arc, or even an atom. If this proposition strikes you as self-evident or even banal, so much the better for psychology. If it does not, at least I have something to talk about.

A generation or two ago it might have been supposed that "physical" meant "material," but in a period when Einstein's equation $E = MC^2$ appears in mass-circulation slick-paper magazines, this is rarely what is meant. When mass is energy and energy may be a local deformation of space, when the particle—is it the pi-meson?—that holds the atomic nucleus together, and thus gives "matter" such permanence as it displays, has itself a duration of a tiny fraction of a millionth of a second, the concept of materiality ceases to be physically meaningful. However, let us call this Meaning 1.

Consider what a psychologist means by "physical" when he talks about a physical (sometimes, to be sure, a physiological) model or a physical correlate. The word seems always, except for an occasional atavistic naivety, intended to mean something like "comprehensible to the physical principles and theory current at the time of writing." Actually, it is more likely to mean "comprehensible to the writer's remembered high school or college physics and haphazard impressions of developments since then." In either event it does not mean the same thing on 13 March 1953 that it meant on 9 August 1935 or on 21 November 1912. And we may be quite sure that it will mean something still different by this time next year. If this is the meaning of "physical" in such expressions, we are not saying anything very definite or precise when we use them. Let us call this Meaning 2.

There must be several psychologists here who remember when it was generally regarded as a complete and final refutation of Thorndike's law of effect to remark that it required a retroactive effect of a later state of affairs upon an earlier and that this was "physically" impossible. Now that the phenomenon of feedback and the construction of servomechanisms have become physical commonplaces, we get along quite comfortably with the notion of reinforcement, as it is now fashionable to call the law of effect. It is even possible now to talk about the self-regulation of organisms and of subsidiary systems thereof without becoming liable to the horrid suspicion of vitalism—provided we call this concept homeostasis.*

We are not yet in psychology far enough from an implied belief in primitive word magic for it to be wholly funny. It is still too stultifying to psychological theory. When we persuade ourselves that it makes us more objective or physiological to call a meaningful object a stimulus, even though it may be what it is for the behaving animal by virtue of a long history of varying kinds of commerce with it, we are indulging in what Harry M. Johnson [*9*] once appropriately called "the method of equivocance." When Pavlov [*14*] called a tissue need or a remote long-term goal a "reflex" he created a confusion that psychology has not yet won clear of. Much as we properly deplore the inherent antinomies and the excesses of certain semantic doctrines, the general injunction to keep our terms clean and their referents unambiguous can be ignored only at the price of talking nonsense.

Let us, at all events, try to get some clarity into the meaning of "physical" as used in psychological discussion. Sometimes the word used is "physiological" or "neurological," but these are clearly regarded as merely special cases or departments of physical. We are told, for example, that unless our hypothetical constructs are in neurological terms they are "forever removed from empirical investigation" [*11*, p. 284]. It may be useful in this connection to consider briefly a series of papers by David Krech, whose egregious and deplorable apostasy [*11*, *12*] from the insights he earlier attained is enough to make strong men weep.†

* Why, by the way, do you suppose that Köhler's admirable 1920 exposition of self-regulation and steady states in physical systems and the eligibility of this concept for psychological use did not suffice to make it respectable for psychologists whereas Cannon's later adoption of it did? Was it merely a happier choice of a name for the phenomenon? Or just that Cannon was duly labeled a physiologist?

† Manifestly I cannot quote extensively enough to supply complete contexts or to be wholly fair. I can only ask Krech to forgive this for the good of our common cause. His reflections are so far in advance of most current psychological thinking that it must seem ungrateful to pick on him. But they are so far from his own high standard that he deserves the worst that compression im-

The general thesis of the second article is that "the moment we introduce hypothetical constructs into our theory building, then the purely psychological approach becomes untenable" [*11*, p. 287]. The clinching argument for Krech, several times repeated, is his question, "Where do these hypothetical constructs exist?" [*11*, p. 284] One is tempted to answer, "In the same place as the binominal theorem or the kinetic theory of gas." The only way I can make sense of this question is to suppose that it means, "Where in space are the data (objects, structures, processes, events, etc.) into which these constructs may be expected, with advancing inquiry, to evolve?" But even this comes very close to implying Meaning 1 of "physical." The application of the seventeenth century criterion of materiality, simple location, is a special case of what Whitehead has called [*16*, p. 85] the fallacy of misplaced concreteness.

To make a long story short, the answer to Krech's crushing question is, "In exactly the same 'physical' world or Nature as the atom or electron." If one has a seventeenth century model of this Nature, or a seventeenth century meaning for "physical," this is bound to be opaque. If, in my answer, "physical" is taken in Meaning 2, it is equivalent to believing that physics is done, finished, exhausted, that there will be no more Plancks, Rutherfords, or Einsteins. Somehow this seems unlikely. We have been hearing a good deal lately about binary calculators, servomechanisms, transistors, and such like. For all we know, or for all physiology or neurology can tell us, our brains or, more generally, our bodies may be and probably are full of transistors, perhaps one in every nerve cell, and a thousand other gadgets that physics has not yet invented, or even dreamt of. When one looks back over the developments of the last 50 years of physics and considers that the rate of discovery appears to be increasing logarithmically, one wonders why we are so timid about our constructs or should insist that they be reducible to the physics of today or even of 50 years hence. Our constructs may be good physics or bad physics, but that does not depend on their intelligibility in terms of present-day physics. It depends rather on their adequacy to the data of behavior—*all* the data, *all* the phenomena, that we have at a given time.

poses upon us. And his dangerous eloquence could easily corrupt the youth of psychology.

There is much, especially in the second paper [*12*], with which one must agree. As a matter of fact, I think his current "dynamic systems" and his earlier "cognitive structures" are one and the same thing, namely sentiments, of which more later.

Any constructs that require us to close our eyes to any of the phenomena of experience are bad constructs and bad physics. I am perfectly willing to "accept the Universe" as Krech [*11*, p. 288] demands, but it does not seem reasonable to reject most of it for the sake of the pip-squeak fraction for which we have at the moment a partial understanding.

It will probably be evident by now that I have smuggled in a third meaning of "physical," one that is neither so patently grotesque as Meaning 1 (materiality) nor so mercurial and stultifying as Meaning 2, with its dependence upon the calendar. In this third sense, "physical" simply means "explanatory." Its opposite is "phenomenal." This meaning was first forced upon me in reading Köhler's *The place of value in a world of facts,* which I had great difficulty in understanding until I read "explanatory" for "physical" in most of its occurrences. Now this Meaning 3 will carry us quite a long way, perhaps as far as we shall get today. At least it may be sufficient to establish the validity of the opening statement of this paper. Its difficulty is the "bifurcation of Nature" that Whitehead [*15*] has indicated one kind of escape from; its implication that cognition is not a natural process but something that is outside of and beyond Nature and mediates the latter to something else called Mind with a capital *M.** This bifurcation is another item of what Krech has somewhere called "Western folklore," in this case Cartesian, in which it has "split asunder what nature had put together" [*10*, p. 82].

When Krech speaks of "a proper respect to present neurological knowledge and theory" [*12*, pp. 345 f.] I think he is dead wrong, in spite of his comprehensive and important qualifications. The only things to which any inquiry owes respect are its phenomena. The attitude of respect on the part of an empirical science is never appropriate toward existing principles of its own or any other field of inquiry. You break out of the bonds of a doctrine and enlarge it only by *not* having respect for it. We are inherently conservative enough without submitting to such restrictions. Thus Planck "five years after Einstein's first publication on the photon theory of light, angrily commented that all the fruits of Maxwell's great work would be lost by accepting a quantization of energy in the wavefront 'for the sake of a few still rather dubious speculations'" [*8*, p. 99]. So far as I know the reconciliation of the wave and the photon theories of light is not yet complete, and you still have to think of light in some con-

* This must not be supposed to depreciate the utility of the really indispensable construct of "a mind" or "a personality."

texts as undulatory and in others as corpuscular. But no one doubts that the integration will somehow be effected. It is an opportunity for some future Maxwell. The point is that you always lose when you ignore phenomena for the sake of principles and may even come out with some such anomaly as a "miniature system," a pedagogical device sometimes useful in teaching mathematics but a self-contradiction in empirical science [2]. The phenomena that Krech seems in danger of ignoring in his retreat from "cognitive structures" to "dynamic systems" are, of course, those of cognition, of which more later.

This is all the more puzzling since later in his second paper [*12*] he welcomes and quotes Bertalanffy's suggestion that physics will enrich and fructify its theorizing through attention to the phenomena of biology. Why does he stop there? Why not recognize (*a*) that physics must ultimately comprehend psychological phenomena and (*b*) that this must come about through the psychologizing of physics rather than through closing our eyes to psychological phenomena? Is this another manifestation of Western folklore, of the Cartesian bifurcation of Nature?

Human cognition is certainly the most highly developed, as well as our most accessible, paradigm for a kind of relation among natural objects. But the fact that physics has not yet got round to dealing with this relation hardly justifies us in ruling it out of Nature or, indeed, in closing our eyes to something rather like it in white rats. I imagine one reason that the term "reinforcement" has so widely replaced "law of effect" is that it permits its users more comfortably to evade this relation. But think how many cognitions are involved in a single reinforcement and try to imagine its occurring without them.

One more quotation from Krech and my consideration of his recent publications will close. He argues that the purely psychological approach

> . . . is untenable because it makes forever impossible any attempt to approach the study of our hypothetical constructs in any more direct manner than through the examination of the original stimulus-response correlations. This is so, I must repeat, because the psychological position places hypothetical constructs in a domain which, *by definition* [italics his], is forever removed from any direct observation (for that domain, it will be remembered, is neither behavioral, experiential or neurological). The conclusion seems inescapable that a psychological field or a life-space cannot itself be a construct nor can it provide us with the substrate in which we can profitably place our hypothetical constructs [*11*, p. 288].

Part of this crushing dictum I have already dealt with: the psychological field is physical in the only meaning of "physical" we have yet found that will make sense. There is a fourth meaning—natural—that also makes sense if you know what you mean by the concept of nature, and in this sense also psychological constructs are just as physical as neurological ones. But we should not overlook the possibility that psychological fields may also be the only fields in nature that are phenomenal. Perhaps psychological field is the only construct in physics that also is directly experienced.

What concerns me here is an epistemological fallacy that has had a paralyzing effect upon American psychological thinking ever since Watson popularized it. This is the widespread belief that our perceptions of "physical" situations and objects are somehow in better epistemological status than our perceptions of other people's psychological situations and objects; that seeing a cat's situation as frustrating is less secure *in epistemological principle* than seeing his physical environment as made of wood or glass. This is simply not so. The only test we have in either case is intersubjectivity. This test may be less frequently or less easily satisfied for psychological situations (although even this is questionable), but when it is satisfied they are just as objective and just as eligible to be causes as are "physical" ones. For psychologists especially it is important to realize that the *epistemology* of values is neither different nor separable from that of things. There is not time to develop this notion here and now. Perhaps its validity is self-illuminating.

But what does Krech mean by "direct observation" in the last quotation? Does he think an atom or an ion or a gene has been "directly" observed? He may have seen a streak in a Wilson cloud chamber that someone told him was the track of a particle of water after impact of an electron. He may have seen a photograph of an image in an electron miscroscope that someone told him was the atomic lattice of a crystal. But these are a long way from "direct" observation. We are apt to forget what an elaborate chain of inference is built into our instruments. And if you are going to worry about transmission theories of light and sound, as a Cartesian will, you must also worry about the transmission of primary qualities. When is any observation direct? And when, above all, are we going to quit wondering whether a given experience is possible and deal with the experience?

I shall not discuss Krech's critique of the hypothesis of psychophysical isomorphism because that can be done more appropriately and effectively by others. My understanding of it differs somewhat from his. I conceive it as primarily a heuristic device, useful in the present state of knowledge, to generate or suggest hypotheses about cortical function which may then be experimentally tested by their implications for perception. Thus, to put it in terms that it is one of the aims of this discussion to transcend, its function is to make not psychology but neurology, and it is of no use to psychologists *qua* psychologists, but only in so far as they turn themselves into neurologists. Actually and historically, of course, it has had extremely valuable psychological by-products, has led to important discoveries that would otherwise have waited a long time.

But I have been much too hard on Krech. Actually, we are indebted to him for creating a climate, providing a basis for communication, that obviates the necessity for this discussion to be even more long-winded than it is. He appears to have been seduced jointly by MacCorquodale and Meehl's "On a distinction between hypothetical constructs and intervening variables" [*13*] and by Hebb's [7] persuasive exploitation of some recent developments in physical understanding. Let us now take a brief look at these two works.

Perhaps critical notice should have been taken of the MacCorquodale and Meehl article when it first appeared. But who could have foreseen that it would mislead a psychologist of Krech's demonstrated quality? And life is much too short, as Wallach once remarked in another context, to spend it rushing about trying to keep all the people who are so inclined from jumping off bridges.

I believe that most of my critique of Krech's position applies with equal or even greater force to that of MacCorquodale and Meehl. In general, their point is that we must give the name of "construct" only to such hypothetical schemes as might "conceivably" be true. This seems reasonable enough until one realizes that their range of conceivability is much narrower even than Krech's. Notions that make no pretensions to represent what they call "objective existence" should be called "intervening variables." These authors are righteously indignant about notions, such as libido, that are introduced as innocent intervening variables and then subtly transformed into constructs by the surreptitious addition of dynamic properties. "These hypothetical constructs, unlike intervening variables are inadmissible * because

* Note the stern and final damnation implied by that word "inadmissible." That really puts them in their place. It might be wondered if the subject of this symposium, personality, is really admissible.

they require the existence of entities and the occurrence of processes which cannot be seriously believed because of other knowledge" [*13*, p. 106].

Don't these people realize that this inadmissible transformation of intervening variables into constructs is precisely the way science is *made?* A little perspective would help here. At what point did Gregor Mendel's intervening variable of alternative hereditary characters become a hypothetical construct? When it was baptized a gene? When reduction division was first observed? Or not until after the electron microscope?

Structural formulae of compounds were at first written as a device to insure that all the "valence bonds" of the elements involved were used. It was really inadmissible later to regard this metaphor as representing the actual structure of molecules. Fortunately the chemists had not read MacCorquodale and Meehl and didn't know this.

It would be interesting, if time permitted us to trace the fascinating development of the notion of the atom, to inquire at what point and by virtue of what inadmissible imputations of properties it became a construct and, indeed, beyond that, how it became a natural object.* Some of you may have seen textbooks printed in this century in which the atom was represented or even pictured as a little round ball with one to four hooks representing its valences, which had to be—inadmissible notion—"satisfied" for the atom to be in a stable condition. The serial modifications of this wretched, metaphorical intervening variable at the hands of Aston, Rutherford, Bohr, and their numerous successors are familiar to you. They have further specified the nature of the hooks, but the hooks are still there.

What I think such a survey would reveal would be something like this: *when a notion shows a good deal of versatility and seems to be applicable to a variety of phenomena beyond that for which it was designed, it becomes a valued construct, irrespective of the immediate plausibility of the mechanisms it envisages.*

At the very end of the MacCorquodale and Meehl article, just before the summary which reiterates its doctrine of sterility, there is a para-

* I called up a physicist friend the other day and asked him, "What is the subject matter of physics?" His response was, "The natural relations among inanimate objects." "And are atoms, electrons, mesons, etc., objects within the meaning you intend?" "Oh, sure!" he said. No epistemological misgivings there.

graph of great wisdom. Unfortunately, both the authors and Krech appear to have missed its significance.* Here is the paragraph:

> Of course this judgment in itself involves a "best guess" about the future. A hypothetical construct which seems inherently metaphorical may involve a set of properties to which hitherto undiscovered characteristics of the nervous system correspond. So long as the propositions about the construct are not stated in the *terms* of the next lower discipline, it is always a possibility that the purely formal or relational content of the construct will find an isomorphism in such characteristics. For scientific theories this is enough, since here, as in physics, the associated mechanical imagery of the theorist is irrelevant. The tentative rejection of libido would then be based upon the belief that no neural process is likely to have the *combination* of formal properties required. Strictly speaking, this is always problematic when the basic science is incomplete [*13*, p. 106].

It ought to be remarked too that Krech's practice of theory construction is far better than his preaching of it. Thus he boldly imputes properties (segregation, which he calls localization, rigidity, etc.) to his construct of dynamic systems for which little plausible basis exists in even the most speculative contemporary neurology, on the excellent ground that, plausible mechanism or no, they *have* to be thus and so to account for the phenomena; just as Mendel's alternative characters *had* to be in the germ cell to account for his data, whether or not reduction division was known to contemporary cytology. Krech may be right or wrong about any of these imputations, but that will be settled by their ability to cope with the psychological data, not by how easily they slip into the strait jacket of contemporary neurology.

But I must also pay my respects to Hebb, our Faust's real Mephistopheles, whose winsome style, wide-ranging scholarship, acute criticism, and candid facing up to the difficulties of his doctrine are an almost overpowering combination. Almost, but not quite. What Hebb has done, very inadequately and roughly described, is to apply some relatively recent notions from physics, notably that of reverberating circuits, to the nervous system and to test the fit for a number of psychological phenomena. Now this sort of thing might be thought to come under the head of good clean fun for psychologists that cannot possibly do any harm and may actually advance neurol-

* They credit Dr. Herbert Feigl with clarification of this point. It is a curious and somewhat depressing commentary on contemporary psychological thinking that two empirical scientists have to have their puristic zeal abated and qualified by a methodologist. But it is very pleasant to be able to say something nice about a logical positivist!

ogy.* But all too frequently, in American psychology at least, the use of the momentarily favored neurological concept comes to be regarded as equivalent to the propositions: (*a*) psychological events not understandable in these terms cannot happen, and (*b*) psychological constructs that cannot be put in these terms cannot be entertained, i.e., are "inadmissible." This happened in the case of the reflex arc, of Sherrington's idea of the final common path, of the all-or-nothing law. It will almost inevitably happen in the case of Hebb's cell assemblies, and what appeared to some as an *extension* of the possibilities of psychological theory construction seems to be already becoming another strait jacket.†

But enough of methodology. It is a sterile and ineffectual business, and I hope that I have said my last word about it. It is emasculating too many of our bright young men.‡ The men who *make* science don't seem to know much about methodology. Newton, for example, said he didn't make hypotheses. It is much more profitable to workers in a young field like psychology to read the classics, the original reports of discoveries in older empirical fields, than to study treatises on scientific method, written by sophisticated contemporary logicians with benefit of hindsight. We have to recognize that not only science but also the scientific method itself is always in the making. The hardest task of the discoverer is always to free himself from the preconceptions he is not aware of having, just because they are part of the intellectual climate of his time, just because they are also taken for granted by the treatises of scientific method. Besides, as Conant has remarked [5, p. 48], *a propos* the extraordinary tenacity of the phlogiston theory, "We can put it down as one of the principles

* Psychologists can have a great deal of this kind of fun making neurological applications of the successive physical discoveries of the next 50 years.

† Even if Hebb and Krech are making good neurology, as I suspect they are, in the same sense that the little hooks on the atom were good physics, the rejection of needed psychological constructs because they do not jibe even with their new and adventurous neurology is still a strait jacket. Why, for example, this fixation on the nervous system? There are a lot of other systems in our bodies, about a few of which we already know a little something. Doubtless the nervous system is some sort of bottleneck in the determination of behavior, but it is hardly the whole story.

‡ Frank Lloyd Wright is credited with the remark that the American culture is the only one that has gone from barbarism to decadence without passing through civilization. This, of course, is merely a libelous witticism as regards the culture as a whole, but it seems fairly to describe the transition of American psychology from Watsonian illiteracies to the contemporary preoccupation with methodology.

learned from the history of science that a theory is only overthrown by a better theory, never merely by contradictory facts." * If methodology really played the role its exponents ascribe to it, *one* contradicted implication of a theory would suffice to dispose of it. But we are all aware of contemporary learning theories that show the same kind of resistance to this supposedly fatal defect· that the phlogiston theory did in Priestley's time.

So let us turn to what I hope I may, with the permission of MacCorquodale and Meehl, call the constructive part of this paper. I have used so much of my time for the critical part that this will have to consist of a series of rather bald assertions on which you are invited to do your own reflecting.

We (and the other animals, at least) live in a world not of stimuli but of objects. I would like to leave it at that; but, because of our unfortunate Cartesian linguistic habits, I have to add that these objects are what they are for a given creature by virtue of the structure of his nervous system and sense organs, the condition of his endocrine system, strength, height, skin color, mood, past commerce with similar objects, and many other things. So I make the temporary concession of calling them *psychological* objects, and their totality at a given time, the animal's *psychological* environment or field. This totality, over longer periods, Lewin has called, following v. Uexküll, the life-space.

All proposed psychological laws appear to be of the form $B = f(P, E)$ or, as some prefer to write it, $R = f(O, S)$, and it is widely believed that all three variables must be physical in Meaning 1 or Meaning 2. But no such law has ever, to my knowledge, been obtained. Probably the most nearly successful such effort is Crozier's [6]

$$\theta = k \Big[\log \sin \alpha\Big]_{\lim 20^\circ}^{\lim 55^\circ}$$

for the orientation of young rats on an inclined plane in the dark. The catch in this is the constant k, which is different for different

* It is too bad that we do not have in America the counterpart of Ostwald's *Klassiker der exakten Wissenschaften.* These are small paper-bound reprints, translated into German when the original language is another, of some hundreds of the real landmarks of science, such as Galileo's *Dialogs,* Claude Bernard's *Experimental medicine,* and Faraday's, Maxwell's, and Gibbs's articles. They are sold for about fifteen cents and are obtainable almost anywhere. There might well be some connection between this fact and the great fertility of German science in fundamental ideas, which Bronfenbrenner [4] has noted in respect of psychology.

species of rat, but which is not independently determined and simply has the value required to make the equation fit the theoretical curve. The guess is hazarded that it *may* represent an unknown complex of anatomical characteristics, such as center of gravity, length of femur, and angle of pelvic girdle. The point is that by this device the physical (Meaning 1 or 2) environment, angle α, is surreptitiously transformed into the psychological environment, or the environment as it is *for the behaving* animal.

No. In order to work, psychological laws have to use psychological variables: i.e., acts rather than responses; organisms as personalities rather than as proton-electron aggregates, pieces of protoplasm, or cell assemblies; and objects rather than stimuli. We have to write the general form of behavioral laws as $B_\psi = f(P_\psi, E_\psi)$. It may happen that B_ψ sometimes coincides with B_ϕ and is partially describable in physical (Meaning 1 or 2) terms, as in the orientation of Crozier's rats, where it can be partially described by angle θ. I say partially, because what the rat is doing is maintaining its balance; and describing this by the aspect that happens to interest the experimenter is like describing the behavior of a dog trailing a fox by saying that he is going north or is moving his legs.*

Now the totality of conditions in an organism that make a given object what it is for the behaving animal, i.e., the general case of Crozier's k, I shall call a sentiment. I have been accustomed systematically to define a sentiment as a part of a personality identified by its reference to an object † and to define personality by pointing, the most fundamental of all operations.

Now this is a concept that, under one name or another, has been found indispensable by nearly all students of personality and by most students of social psychology. It has been variously called disposition, mental system, psychical system, means-end readiness, attitude, belief, derivation, metanerg. Krech used to call it cognitive structure but now regards the notion as illegitimate. He has imputed some of its properties to his present dynamic systems, and, knowing how re-

* Zener has pointed out how thinking about the conditioned response is vitiated by restricting consideration to one aspect of the animal's behavior [*17*].

† This includes, of course, the corresponding action system: a thing is for us what we do with or about it. There is a very good idea at the center of the old motor theory of meaning. All it wants is a little cleaning up and moderation of its claim to be the whole story.

It also includes the affective component, an aspect of which I have recently dealt with [3]: i.e., an object is also what we feel about it, as well as what we know and do about it.

sponsible he is to the phenomena, I feel safe in predicting that he will soon impute the rest of them.

Now if the concept of sentiment should prove useful in the understanding of learning it would have demonstrated some of the versatility that I suggested earlier is more important than ease of translation into contemporary physical (Meaning 2), physiological, or neurological terms, and, incidentally, it would further the synthesis with which this symposium is concerned.

I believe that the same assumptions that are made for purposes of personality theory and perception suffice to further somewhat our understanding of learning.

It would seem natural to suppose that my sentiment for (attitude toward, cognitive structure for) this building is in some sense a part of my sentiment for the class buildings. To put it more generally, if object A is a member of object B, sentiment A' is a member of (or part of, or region of) sentiment B'. Moreover, if object A is a *property* (or a *sign*) of object B, it seems evident that sentiment A' must in some sense be a member of sentiment B'. Thus, my sentiment for the height of this room is a part of my sentiment for this room. In general, it would seem that *all* the constitutive relations obtaining among a personality's objects can be ordered to relations of membership among the sentiments corresponding to those objects.

Of peculiar importance among these is the relation of instrumentality: if A is instrumental to B, positively or negatively (i.e., as means or as barrier), sentiment A' is a member of sentiment B'. If we use the sign $\rightarrow$ to signify "is a means to" and the sign $\subset$ to mean "is a member of," the notion could be expressed by the equivalence:

$$A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \cdots \equiv A' \subset B' \subset C'' \subset D' \subset E' \cdots$$

Thus, a given hierarchy of means-end relations among the objects in a person's psychological environment is represented in the personality by a sentiment with precisely the same number of hierarchically organized member sentiments on the same number of membership levels. If all the objects of a given person's environment are instrumentally related to one, then all his sentiments are members of the sentiment for that object and the personality is said to be integrated [3].

Consider the naïve dog in the classical Pavlov experiment. He is alone in a strange and silent room. For a more or less protracted period nothing happens except the events of his own body. Then suddenly there is a buzzing noise off to his right. He has a sentiment for noises. He pricks up his ears and looks in that direction. His back hair may rise. Before he has done reacting to this event, another occurs: a biscuit rattles down a tube and appears in a pan immediately in front of him. He has a sentiment for biscuits and knows how to deal with them.

Now look at the same dog after ten or twenty such sequences of events. Now, when the buzzer goes off, he does not look in that direction but lowers his nose to the pan and gets the biscuit before it has stopped rolling. The buzzer is now another object than it was at first: it is a sign or property of the biscuit in much the same way as the taste, smell, or hardness, which had also at some time to be learned. And according to our assumption about instrumentality the dog's buzzer sentiment has been incorporated in his biscuit sentiment. How such irrelevant things as a buzzer and a biscuit can be integrated into one object I have described elsewhere [*1*, pp. 170 ff.].

This change in the relations or "meaning" of an object seems to be the one objective thing that is common to all the processes that have been called learning. Baptizing a concept does not ordinarily accomplish much, but occasionally it does reveal an identity that might otherwise escape notice. If, as I believe, the conditions within an organism that make an object what it is for that organism—the general case of Crozier's *k*—can be identified with the concept of sentiment that personality theory finds indispensable, a solid basis exists for the integration of learning theory and personality theory.*

REFERENCES

1. Adams, D. K., A restatement of the problem of learning, *Brit. J. Psychol.*, 1931, *22*, 150–178.
2. Adams, D. K., Note on method, *Psychol. Rev.*, 1937, *44*, 212–218.
3. Adams, D. K., The organs of perception: sentiments, *J. Pers.*, 1953, *22*, 52–59.
4. Bronfenbrenner, U., Toward an integrated theory of personality, in Blake, R. R., and Ramsey, G. V. (Ed.), *Perception: an approach to personality*, New York, Ronald Press, 1951, pp. 206–257.
5. Conant, J. B., *On understanding science*, New York, New American Library, 1951.
6. Crozier, W. J., The study of living organisms, in Murchison, C. (Ed.), *The foundations of experimental psychology*, Worcester, Clark University Press, 1929, 45–127.
7. Hebb, D. O., *The organization of behavior*, New York, John Wiley, 1949.
8. Holton, G., On the duality and growth of physical science, *Amer. Scientist*, 1953, *43*, 89–99.
9. Johnson, H. M., Some fallacies underlying the use of psychological "tests," *Psychol. Rev.*, 1928, *35*, 328–337.

* Professor Krech, who was good enough to read this paper soon after the Symposium, thinks my selection of quotations seriously misrepresents his views. Since he may be right and since the articles in question are exceedingly valuable in any case, I would urge the interested student and theorist to read them more carefully than Krech thinks I have.

10. Krech, D., Notes toward a psychological theory, *J. Pers.*, 1949, *18*, 66–87.
11. Krech, D., Dynamic systems, psychological fields, and hypothetical constructs, *Psychol. Rev.*, 1950, *57*, 283–290.
12. Krech, D., Dynamic systems as open neurological systems, *Psychol. Rev.*, 1950, *57*, 345–361.
13. MacCorquodale, K., and Meehl, P. E., On a distinction between hypothetical constructs and intervening variables, *Psychol. Rev.*, 1948, *55*, 95–107.
14. Pavlov, I., *Conditioned reflexes*, Oxford, 1927.
15. Whitehead, A. N., *The concept of nature*, Cambridge, 1920.
16. Whitehead, A. N., *Science and the modern world*, New York, Macmillan, 1925.
17. Zener, K. E., The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response, *Amer. J. Psychol.*, 1937, *50*, 384–403.

Ego Psychology, Cybernetics, and Learning Theory

O. H. MOWRER

It will best serve our present purposes if I discuss the three topics constituting the title of this paper in the reverse order of that in which they are here mentioned. But first a more general word. It can hardly escape even commonplace observation that we tend to take, as models for interpreting the complex and mysterious, phenomena which are simpler and more fully understood. Hence, the machine, being man-made and intelligible, has often patterned our thinking about the less intelligible aspects of man himself.

Today we face a new challenge in this respect. Stanton and Sylva Cohn, writing in a current issue of *The Scientific Monthly,* put the matter well when they say:

> The nineteenth century was the "Age of Power." It saw the development of the machine, and concomitant with it there arose a mechanistic philosophy of life and a mechanical interpretation of life processes. . . .
>
> Science has advanced beyond the mechanistic stage, however. Just as the nineteenth century was the Age of Power, the twentieth century is the Age of Communication and Control. It is not enough to make a powerful machine, having the ability to do many times the work of man. There must be an intelligent application of this energy—it must be controlled [*1*, p. 87].

In recent decades engineering has moved rapidly forward along these lines, producing, oddly enough, machines that are more "intelligent" in practice than living organisms are in theory! Of course, some of these machines are *actually* more "intelligent," as regards certain specialized tasks, than are animals, including men. Here we think particularly of the "giant computers," for example. But we are presently concerned rather with the extent to which living creatures are, per hypothesis, more limited in their potentialities than we know them, in fact, to be. Modern machines are thus presenting a challenge to our theorizing in psychology and related sciences. If we can meet this challenge, psychology may, as Harry Harlow hopefully opined a few years ago, "eventually catch up with common sense." It is with this challenge and some of the new vistas it opens up that this paper will be mainly concerned.

I. THE PASSING OF "HABIT" AND THE REDISCOVERY OF "CONSCIOUSNESS"

In the paper already cited, Cohn and Cohn say:

> If there is one law that marks this era [the nineteenth century Age of Power] definitively, it is the principle of the conservation of energy. This principle, which is expressed in the first law of thermodynamics, has been characterized as the greatest generalization in natural science. But it is not the final word [p. 87].

"Habit" is a concept born of this tradition. A stimulus, as *cause*, impinges upon a sense organ (internal or external) and sets up neural impulses which, by virtue of certain neural "connections," travel to and activate certain muscles. The resulting response is the *effect*. Energy, though transmitted and transformed, is thus strictly "conserved," in the manner of a moving billiard ball striking and imparting its momentum to a second ball, it to a third, and so on. But thus far we have not distinguished between "habit" and "reflex." Reflex, we are told, is invariable, unmodifiable; habit, on the contrary, can be changed. But how? Thorndike noted that, with a habit, the cause-effect sequence does not end with response. Responses, he observed, may in turn initiate causal sequences in the external world which terminally impinge back upon the organism. These "feedback" effects Thorndike, like the layman, called rewards if they lessen stimulation in some important way, and punishments if they significantly increase stimulation. Rewards, Thorndike conjectured, strengthen the *S–R* sequences that produce them, whereas punishments have the reverse effect.

That an organism that can be thus modified by experience—that can, in other words, *learn*—will on the average be better off than a purely reflex organism is pretty obvious. But the model or image which Thorndike gave us has not been a universally satisfying one. On the one hand it has been charged, rather unjustifiably it seems, with being "teleological" (in the opprobrious sense of the term); it has also, more relevantly, been accused of making organisms more "mechanical," "blind," "stupid" than they really are. Although certainly more "intelligent" than a purely reflex creature, Thorndike's "habit" animal is by no means overburdened with brightness. Yet it has not been too easy to say exactly what is wrong with such a creature and how it might be improved.

Here we will attempt a concise diagnosis. It follows from what has been said that if a Thorndikian animal has acquired, under one set of conditions, a given "habit" and if conditions are now *changed*, the animal itself can begin to change *only after* it has performed the old response under the new conditions *at least once*. In other words, this much stupidity, or "maladaptation," is logically demanded by the theory, and has been rationalized by the slogan: "Organisms learn only by doing!" Evidence recently reviewed in another paper [7] indicates that this inference is plainly not valid and calls for a radical revision of what we have previously called learning theory.

In order to escape from the Thorndikian dilemma, we must, first of all, repudiate one of his major assumptions: we must abandon the idea that rewards strengthen stimulus-response bonds and that punishments weaken them. What they do instead is to produce, by conditioning, *secondary reinforcements* and *secondary motivations*, respectively. This will at first hardly seem like a clarifying statement. What it means, quite simply, is that we do not learn, or fixate, overt, behavioral responses at all. These are always "subject to change," depending upon the "situation." What *is* learned are attitudes, meanings, or expectations which consist of token decrements in emotional tension (secondary reinforcements, or rewards) and token increments (secondary motivation, or punishment). It is assumed that it is these inner, conscious factors which, moment by moment, select and shape overt action; and if we take this position we have ample provision for "learning" *without* doing, i.e., for changes in behavior that occur, solely and immediately, because the *situation*, or, more exactly, the individual's internal tension state, or "field," has changed. Here we have the capacity for foresight, insight, and a generally higher order of intelligence and adaptivity than is possible in a "creature of habit."

But we have achieved this at a cost which some will be reluctant to pay. Instead of channeling stimulus energies directly through the nervous system, switchboard fashion, and out into motor organs in a highly determined way, we are here assuming that this kind of determinism holds, so to say, only half way. It holds, I assume, to this extent, that meanings and attitudes, both positive (tension reducing) and negative (tension inducing), follow quite automatically, quite reflexly (conditioned reflexly), upon the occurrence of significant stimuli or situations. But here this type of fixed, cause-effect relationship ends and a more complex mechanism takes over. I am sure that you will not hold me accountable for explaining all the riddles of consciousness merely because I refer to the phenomenon; but I will venture the guess that consciousness is, essentially, a *continuous-computing* device or process. The eternal question is, "What to do? How to act?" And consciousness, as I conceive it, is the operation whereby information is continuously received, evaluated, and summarized in the form of "decisions," "choices," "intentions" which then emerge as behavior. Life asks the questions, sets the problems, and it is the business of consciousness to give us the "answers." *

This is not to say that consciousness is merely chance or caprice. In the paper already cited [7], an attempt is made to state some of the principles of conscious functioning, and more attention will be given to this problem in the later sections of the present paper. I hope that I have succeeded, thus far, in showing the general direction in which learning theory must, in my judgment, move if our conceptual models in psychology are to be as resourceful and sensible as the "real thing" or even as the newer types of machines that can today do such remarkable things.

II. CYBERNETICS AS THE SCIENCE OF COMMUNICATION AND CONTROL

It will, I trust, be evident how naturally the foregoing analysis articulates with some of the basic concepts of cybernetics. One of these concepts, as Wiener [*10*] and others have formulated it, is that only the simpler machines and response systems operate on a "blind," reflexive principle; "higher" behavior and machines involve the "feedback" principle. Learning, I propose, is not a matter of strengthening or weakening connections between drives and overt behavior, but of

* The reintroduction of the concept of consciousness into behavior theory may seem, to some, like a retrogressive step. To such persons it may connote a return to introspection as the chief mode of psychological inquiry and a relinquishment of the gains of half a century of intensive experimental inquiry. What I wish to suggest, without being able to develop at this time, is the notion that consciousness is a phenomenon which would be inferable from, and indeed logically demanded by, the empirical facts, even though there were no direct experiential access to it whatever. The point is that adherence to a strict S–R psychology creates a number of dilemmas which only the interpolation of an *integrative* mechanism of some sort can resolve. Consciousness is here conceived as that mechanism.

the acquisition of "positive" (rewarding) and "negative" (punishing) feedback from stimuli that have accompanied past action or experience. If, in the past, a given act has been predominantly rewarding, then incidental stimuli, both external and internal, which have been associated with the act-reward sequence will take on, as already noted, the capacity to produce secondary reward and thus to guide or direct the organism into the *same* or similar action in the future. If, on the other hand, a given act has been predominantly punishing, the incidental stimuli which have been associated with the act-punishment sequence will take on the capacity to produce a secondary-drive increment ("fear") and will tend to guide the organism into *different* behavior.* Behavior, in any given situation, is the organism's best effort, then and there, to find that line of action with the greatest positive and the least negative feedback—and hence with the best likelihood, when it is carried through to completion, of being maximally satisfying and minimally hurtful. Feedback has been called "the central theme of cybernetics." We must, it appears, likewise give it a major role in psychological theory.

Another way of saying much the same thing is to stress the role of communication and control. These are the factors which are conspicuously missing in most nineteenth century and earlier machines and which are also minimal or absent in reflex action in living organisms. Central to the communication process, surely, is the phenomenon of *meaning* [8], and this we have given a central place in our theoretical scheme. Meanings, not means; attitudes, not actions —these are the most immediate outcomes of learning; and it is through these that intelligent *control* of behavior then becomes possible. It is not yet certain whether so-called "information theory" as it has been elaborated by Shannon and Weaver [9] and others will prove as helpful in behavior analysis as many psychologists now hope; but it seems abundantly clear that the more rudimentary tenets of cybernetics have already given our thinking in the realm of learning theory a useful nudge; and I shall try to show in the next and last section of this paper that the developments thus produced carry us appreciably closer, as learning theorists, to ego psychology than we have been before.

First, however, it will be salutary to clear up a terminological point. In the preceding pages I have spoken of positive and of negative feedbacks, defining them as secondary reward and secondary punishment,

* These principles are here stated in highly condensed form. For elaboration and exemplification, the reader is referred to the longer paper already cited [7].

respectively. Here it should be said that for Wiener and other cyberneticists these terms have a different meaning. If a control system, like a thermostat or the governor on a steam engine, has a *stabilizing* influence on some quality—heat, in the one case; speed, in the other—then the feedback is said to be *negative*, regardless of whether it is acting at any given moment to bring a quantity *up* to or *down* to a standard state. The "error," in other words, which the control system is trying to "correct" may be either "plus" or "minus," but the type of over-all control is, in either case, said to involve negative feedback.

The opposite, or *positive*, type of feedback is no less interesting. The products of certain chemical reactions are such as to have a catalytic effect upon and thus to *speed up* the reactions which produce them. Sometimes these reactions gain momentum so rapidly as to produce "explosions"—they are, in fact, the same in principle as the chain reactions in the atom bomb and the H-bomb. If, likewise, certain other types of processes start to slow down, the effects of the retardation accelerate the slowing down. Thus, if a business concern is not prospering, inability to pay employees satisfactory wages or to meet creditors may accelerate the failure.

Since homeostasis, or self-regulation, is one of the essential characteristics of living organisms, it will be immediately evident that many of their activities will have the properties of negative feedback, as just defined. These activities or functions serve to hold certain states or qualities within fairly narrow limits of variation. Positive feedback, when the concept is applied to living organisms, sounds as if it would be pathological, to say the least, and, in the extreme case, lethal. Actually, we find some very instructive instances of it in living organisms. We have already seen that there is a general tendency for incidental stimuli which have been associated with primary-drive reduction to produce secondary-drive reduction and for such stimuli as have been associated with primary-drive induction to produce secondary-drive induction. These two types of contiguity learning, or conditioning, unquestionably tend to produce (indirectly, through the integrating function we call consciousness) responses which have biological utility—notably, flight or immobility in the face of danger and approach to objects or situations with rewarding potentialities.

However, there is one crucial respect in which this general scheme is inadequate, non-biological. The world's good things do not always remain conveniently at rest, waiting to be claimed and consumed. Sometimes these goal objects, especially when they are other organisms, have a way, at the critical moment, of eluding their pursuers. Therefore, an organism that became more and more confident and *relaxed* as it approached a quarry might find itself slowing up at just the point when a final "push" was needed for success. It is therefore interesting and altogether understandable that we should find the phenomenon of *appetite*. Its outstanding feature is an *increase* in secondary motivation just as consummation is imminent, thus giving to behavior at this crucial moment a peculiar urgency and "oomph."

Here, it seems, is an instance of *positive* feedback: "The nearer you get, the more you want it" carries the proper connotation here. Certainly such an arrangement is biologically intelligible, but it is nonetheless enigmatic. Neal Miller [3] has put the matter this way. We have reason for thinking that stimuli associated with consummatory states take on contradictory capacities: a tendency to cause a decrement in secondary drive (secondary reinforcement) *and* a tendency to cause an increment in secondary drive (which we call, not punishment, but appetite). As Miller points out, if these two tendencies occurred simultaneously, they would be self-canceling, mutually neutralizing; so his proposal is that they may alternate, producing intermittent bubbles of pleasurable anticipation and surges of intensified drive. This is frankly a speculation and, even if true, leaves many unanswered questions. However, consideration of the problem in the cybernetics setting illuminates and sharpens it in ways which we shall further explore in the next section.

III. PSYCHOLOGY OF THE EGO AND SUPEREGO

It is at once evident that neither reflexology nor habit theory, in the mode of Thorndike, Pavlov, Watson, or Hull, can provide a very sophisticated approach to ego psychology, so-called. The reflex and habit, almost by definition, exclude consciousness, which is the core of ego functioning. The same difficulty does not at all occur in the conceptual framework presented in the preceding sections of this paper. Here we posit a division of labor between learning as a purely unconscious, automatic process, on the one hand, and conscious judging, deciding, and acting, on the other. Our conception of learning therefore blends naturally enough with ego theory.

Thinking, in this frame of reference, we see as a form of activity in which the individual makes symbolic responses (Hull's "pure stimulus acts") for the purpose of eliciting, or "sampling," the feedback, or effects, that could be expected if the action thus symbolized or represented were really carried out. In simplest, most primitive form, thinking, or reasoning, can be *seen,* quite literally, in the vicarious trial-and-error of a rat at a maze choice point—a form of "light" activity which is manifestly a prelude to grosser and more consequential action. In adult human beings the process becomes both more subtle and more elaborated, but its function, basically and ultimately, remains the same.

Recently I heard a college president say that the most important thing his institution could do for students was to "teach them to think." At one level this can be regarded merely as an old pedagogic bromide. Or it can be seen as profoundly and perennially true. If in thinking we are, so to say. sticking our "mental necks" out *into the future* and "feeling around," nothing, at least on occasion, could be more useful, especially when the findings are later translated back into intelligent action. Living organisms swim forward in a sea of time, and those with the best "distance receptors," i.e., with the best symbolic skills, will almost certainly have an edge in the struggle for existence. "Use your head" (or head end) is good advice both in the sense of using the special senses and in the sense of moving back and forth through time in the way that symbols make dramatically possible.

Then, too, our conceptual scheme puts us in a good position for understanding not only thought but also *fantasy.* The difference, basically, is that thought is a *preparation* for action; fantasy is a *substitute* for it. In fantasy we select not those symbols that will forecast reality but rather those that will yield, without reference to reality, the greatest present pleasure, or secondary reinforcement. "In daydreams I often picture myself as a very generous and kindly person," a young woman undergoing psychotherapy recently remarked. This was said in the context of revealing just the reverse real characteristics and shows the highly autistic nature of such activity.*

So far, however, we do not yet have a psychology of the abnormal, a psychopathology. Does our system, as presently conceived, yield one? The attempts of reflexologists and habit theorists to "explain" neurosis and propound a therapy have not been very adequate [6]. A minimal essential for a psychopathology is *conflict,* a concept which is hardly meaningful unless consciousness is posited. But conflict is not a *sufficient* cause of neurosis. Indeed, if our view be correct, consciousness is a continual interplay of contending forces;

* Another patient, a young physicist, reports excessive reading of science fiction. "In *it,*" he observes, "the experiments *always* come out right!"

and decision making, compromise, and integration are its major accomplishments.

Where, then, does pathology arise? Section II of this paper gives us a clue. There we have seen that there are two broadly different principles of feedback. If feedback functions so as to make an organism move faster when it lags or slow down when it is going too fast, it is said, by the cyberneticists, to be a negative feedback; and we are at once reminded of the role of conscience or superego. This agency within the total personality is said to be wise, prudent, *balanced,* and mainly concerned with (social) *norms.* The mentally aberrant are said, per contra, to be unstable, unbalanced, abnormal. How do they get that way?

Through parental training as well as by direct experience, human beings acquire conscience, which, as we have seen, operates on the negative-feedback principle. Other things equal, this should be compatible with the over-all tendency toward self-regulation and homeostasis. However, we have already noted a complication. Man has appetites and lusts as well as conscience; and if one operates on a conservative, judicious principle, the other is prodigal and reckless. Sex is powerfully appetitive, clearly following a positive-feedback principle of the "explosive" variety, and anger is likewise the occasion for our sometimes "blowing up." Little wonder, then, that these human proclivities are the most difficult to "control" and the ones which most often come into conflict with superego functions. Caught between two such powerful forces as id and superego, appetite and conscience, lust and guilt, what indeed is the "answer"? Here the integrative resources of the ego are likely to be taxed to their utmost!

The Freudian version of the etiology of neurosis holds that in the struggle between negative feedback (conscience) and positive feedback (appetite), the former sometimes seizes too firm a control over the latter and thus holds down, or inhibits, "instinctual" functioning to a pathogenic degree. Anxiety and depression are said to be the fruits of this stifling dominion of the superego over id functions [2].

On other occasions [*4, 5, 6*], I have argued for the contrary view, that neurosis arises when positive-feedback functions have out-contended the negative feedback, resulting in temporarily uncontrolled, "explosive" behavior for which the individual later feels remorse and shame (intensified negative feedback). The ego may then deal with this expression of an "aggrieved" conscience in several different ways, notably by making a confession and amendments, in which case inner stability and harmony tend to be restored; or the ego may set out to "disconnect" the conscience, like the governor on a steam engine, so as to let the "wild" behavior occur unobstructedly (albeit usually surreptitiously). Conscience may be thus dissociated or repressed, and the individual may compliment himself on his new freedom, liberty, emancipation. But the forces of conscience are tenacious and usually come back to haunt their owner, but now not as intelligible guilt but as the unintelligible and torturing experiences of anxiety, panic, depression, and inferiority feeling.

It is not my wish or purpose here to debate the relative merits of these two interpretations. I present them rather to show how neatly and naturally they can both be accommodated in the general theoretical framework here presented, one in which the view of "habit" as *S–R* bonds is rejected in favor of a more complicated system in which learning is limited to the acquisition of inner meanings which are then evaluated and used to arrive at the particular (often novel) decision and action which the total situation seems most to warrant. One can hardly escape the feeling that cybernetics contains some powerfully unifying principles and that science may yet lead to a type of synthesis of human experience which has, by some, long been held impossible.

REFERENCES

1. Cohn, S. H., and Cohn, Sylva M., The role of cybernetics in physiology, *Scientific Monthly*, 1953, *76*, 85–89.
2. Freud, S., *A general introduction to psychoanalysis*, New York, Liveright Publishing Corp., 1935.
3. Miller, N. E., Learnable drives and rewards, in Stevens, S. S. (Ed.), *Handbook of experimental psychology*, New York, John Wiley, 1951.
4. Mowrer, O. H., *Learning theory and personality dynamics*, New York, Ronald Press, 1950.
5. Mowrer, O. H., *Psychotherapy–theory and research*, New York, Ronald Press, 1953.
6. Mowrer, O. H., Neurosis: a disorder of conditioning or problem solving? in Miner and Kemp (Eds.), *Comparative conditioned neuroses*, New York Academy of Sciences, 1953, pp. 273–288.
7. Mowrer, O. H., Is "habit" merely "secondary reinforcement"? (unpublished).
8. Osgood, C. E., The nature and measurement of meaning, *Psychol. Bull.*, 1952, *49*, 197–237.
9. Shannon, C. E., and Weaver, W., *The mathematical theory of communication*, Urbana, University of Illinois Press, 1949.
10. Wiener, N., *Cybernetics*, New York, John Wiley, 1948.

Personality Structures as Learning and Motivation Patterns—A Theme for the Integration of Methodologies

RAYMOND B. CATTELL

We have come together at this symposium hoping for an easy and happy marriage of two young branches of psychological science; but the tragedy of Romeo and Juliet should remind us that the failure to consider rivalries of parentage can lead to difficulties even when true love exists between the parties—and I am not at all sure that it does in this case! Consequently if our purpose is to arrive at joint concepts and methods we must first face some embarrassing inquiries about the implicit viewpoints and ancestries of the parties concerned. As a personality theorist, my role at this marriage is perhaps that of mother-in-law to learning theory, so your expectations of sweet reasonableness on my part must be according to your personal projections in this situation.

Quite apart from these immediate differences we have to recognize that our more remote common ancestor, psychology, was itself a problem child among the sciences, always boasting to answer the spectacular questions—and by polysyllabic global theories—before it acquired the patience and method effectively to answer the small questions about behavior. Theoretically, for example, it indulged in such pretentious mathematical-sounding names as topology and in such elaborately complete motivational and structural systems as those of psychoanalysis,. without first establishing accurate methods of description and measurement upon which reliable and worthwhile laws could be based, or determining personality structure by objective, multivariate, factor-analytic methods. Second, it failed to develop objective test measures for whatever patterns had been shown to be functionally independent or dimensionally clear. When this is done, and only when, is it economical to begin investigating by learning theory the etiology of personality structure. Much valuable research effort would be wasted if exact workers in learning theory were persuaded to explain the development of normal and abnormal personality formations as they exist at present in the shaky hunches and formulations of clinicians. For though the clinic is admittedly one of the best places in which to study personality the *implicit* multivariate analysis which goes on in the clinician's head, when he uses the global, wholistic arts of the clinical method, has surely to give way to the more *explicit* multivariate approach of advanced statistics. Thus only can the results emerge undistorted by personal prejudice and be subject to verification by precise hypothesis testing.

The personality theorist's position is in many ways intermediate between that of the clinician and the learning theorist, in respect to interests, methods, and theoretical tools. Accordingly, it is up to him to indicate, since he stands on the ground where the integration is most likely to occur, what difficulties he has in accepting within a common system some of the proposals from either flank.

It is comparatively easy to indicate those encumbrances which I think we should *not* take over from present clinical research. The main thing we should not take on is the vast proliferation of mere speculations, now widely accepted as truths, which occurred before the prodigal psychologist repented of his prodigality and recognized the necessity of a planned sequence of objective research such as I have described. This prodigal phase has been costly enough in research time even without our also adopting the debt of its loose conceptual offspring. It has filled our clinics with trashy, patent-medicine measures of personality, it has littered our journals with researches upon the slippery foundations of which more exact scientists cannot build, and it has filled the minds of our students with glib, ready-made explanations where there should be honest doubts and clear perceptions of issues open to research. Worst of all it has almost totally filled clinical positions with a generation which is largely innocent of scientific method, which has its interests heavily vested in particular crystal balls and test gadgets, and which has no notion, for instance, that a criterion is necessary to determine if therapy works at all. Cloistered researchers may say that they meet none of these in scientific discussion, but they are at least *social* realities.

Essentially, however, we have turned the corner from this phase, thanks to such rigorous measures as the A.P.A.'s definition of a clinician as a fully trained psychologist, and I mention this sad background only because we have to distinguish in contemporary discussion of clinical contributions between what is obviously *socially* accepted and respectable and what can be considered *scientifically* respectable. I shall, therefore, take the position that clinical psychology has contributed and will continue to contribute a lot of valuable hunches about personality for the process of scientific testing. On the other hand I cannot take the position that some of my clinical friends would like, to the effect that the experimentalist is someone who merely checks on clinical hypotheses. For increasingly the valuable hypotheses in personality research are arising from experimental work itself or at any rate from factor-analytic work.

The difficulties on the other flank, in relation to learning theory, are those between people with similar technical standards but different histories. Yet there is one simple respect in which it should be said initially that learning theory is more seriously disabled than clinical psychology, namely, that whereas clinical psychology has had extensive empirical contact with personality problems, learning theory has had virtually none. Since I have a strong predilection for laws that develop out of the phenomena to which they are supposed to be relevant, I cannot feel confidence, scientifically, in learning laws taken over ready-made from another area, unless and until they justify and enlarge themselves in relation to personality measures themselves.

The proposed ideal research strategy of first determining personality structure and then determining its laws is therefore embarrassed from the beginning by attitudes and conditions in the history of psychology and psychologists, and in this case I favor Oscar Wilde's definition of history, namely, "an account of things that should never have happened." However, if we may aim even now to institute a more ideal and effective sequence I suggest that we recognize five steps, or phases, and I propose to list these and devote 10 minutes' discussion to them. First there is a phase of precise, quantitative investigation of existing structure in the organism, in this case of human personality structure [2]; second there is the step of producing reliable instruments for unitary measures of these meaningful structures and functions [5]; third—and here I ask you to bear with my eccentricities—I would ask for an initial investigation of the part played by constitution and heredity in these measured patterns and structures, so that we do not waste our time trying to explain by learning theory what is really a maturational phenomenon. The fourth stage would be the application of the usual stimulus-response learning experiment design involving sequential, causal, longitudinal analysis of what has first been studied in cross-section, in phases 1 and 2. I say the usual learning experiments, but actually this etiological investigation would also go beyond controlled experiment. I urge, in addition, new multivariate designs for studies *in situ,* by methods I hope to indicate later. Fifth and finally we can expect a phase in which laws of learning have themselves become enriched and perhaps modified by the new problems and new kinds of material encountered in personality change, by a feedback of these augmented principles into the laboratory.

Perhaps 10 years hence some psychologist will be able to illustrate by pointed examples the principles encountered at each of these phases, but my own expansions of them are today very patchy. Researches directed to even the first two phases have been so little and so late that at present one has to turn to articles on which the ink is scarcely dry in order to illustrate steps penetrating no further than the third phase, and beyond that lies only speculation or work lacking the foundations indicated. Indeed, as far as I know there are just four studies in existence that break into the third phase, that of separating maturational from learning patterns—namely, Eysenck's [*10*] studies on alleged neurotic patterns, Thurstone's work now in progress on primary abilities, a study at Michigan, and the Illinois study on personality factors in twins [*7*]. In this Illinois study we are measuring about a dozen known personality factors by objective tests on about a thousand identical and fraternal twins, siblings reared apart, unrelated children reared together, etc., in order to determine nature-nurture ratios for major source traits, as was done some years ago for the single source trait of general ability. From this we hope a number of definitely learned patterns will clearly segregate to be submitted to learning experiment. However, even without the possibility of exemplification in phases 4 and 5 it will help discussion of the problems therein if we pause to glance more closely at the nature of some of the patterns found in phases 1 and 2 [*2*], for in these factors reside the concepts we have to work upon with learning theory. Operationally the basis of these patterns is an *R*-technique factor analysis [*1*] of variables carefully chosen to cover the personality sphere, with subsequent discovery of a unique factorial simple structure, obtained "blindly" in order to be uninfluenced by the experimenter's theoretical prejudices [*6*]. In this glance at the patterns in question I shall not take stock of life-record, clinical-rating data or questionnaires, which are properly only preludes to factors in objective, situational tests. At least I shall not do so except in so far as the behavior rating factors permit a matching of the objective-test patterns with clinical and general criteria.

It is well known that factors have been obtained, and confirmed on new populations, which correspond to the cyclothyme-schizothyme dimension of the clinicians, to the concept of emotional maturity or ego strength, to the structure of the superego, to general anxiety level, and to some further six or more dimensions which differ from those just named in that they transcend present clinical insights [*2*]. Though it is encouraging for hopes of integration of clinical and quantitative methods to find some degree of convergence of concepts it is also not surprising to a methodologist that these finer measurement and analysis methods yield, as indicated, newer patterns unknown to the clinician, just as the microscope yielded species once unknown to the biologist. Moreover, although the clinical and factor-analytic findings undoubtedly tend to converge, the latter point also to important modifications required in the clinical formulations.

For example, the factor analyses clearly show two factors, *A* and *H*, not one, in the area of schizothyme behavior, and one of these, *H*, has physiological, autonomic associations, suggesting some constitutional element, whereas the other, *A*, does not, but has patterns of hostility and rigidity suggesting the result of general environmental frustrations. Further, one can find a factor of paranoid trends, *L*, quite distinct from the schizothyme factors as such, which finding clarifies certain clinical classification obscurities [*2*]. Another instance where the clinical picture is both clarified and modified by the factor analysis and its reference to normal populations is found in the factor which loads the variables concerned with conscientiousness, perseverance in maintaining standards, altruistic regard for other people, and other manifestations identifying it with the superego conception. This factor is virtually uncorrelated with factors of general neuroticism (as Mowrer's hypothesis [*15*] would also require), and its pattern is not overly loaded in guilt feelings, which suggests that the clinician's view of the superego gained from neurosis formation has unknowingly been strongly colored by the biased population sample taken.

As I have already indicated, the methods of the clinician and the factor analyst, being wholistic and multivariate, are in a larger sense the same, but the clinician uses the analytical powers of his own memory whereas the factor analyst prefers the less subjective and more refined instruments of quantitative records and calculating machines. The important thing initially is that their results agree. But it remains for further research, particularly the pioneering of such factor analysts as Ferguson, Wittenborn, and Eysenck [*10*], concentrating on abnormal populations, to establish to what extent such differences as I have indicated between factor analyses of the normal personality structure and clinical views gathered from the abnormal arise, respectively, from differences in objectivity of techniques, and from the contrasts of normal and abnormal population samples.

When now I turn, as promised, to factors in *objective* personality tests themselves I am compelled, in contrast to the discussion of rating factors, to ask you to take on trust a weight of evidence which, like the bulk of an iceberg, is still largely invisible (and unpublished). Since 1944 some seven independent studies, related in a progressive sequence converging on gradually clarifying patterns, have been carried out in our laboratory alone, and only two have yet been published. They cover over 200 different types of personality tests and a population of nearly 2000 cases. The upshot is that some seven or eight factor patterns have now recurred so consistently as to give us considerable confidence in their being very stable source traits in personality.

By way of illustration we may take that factor [5] which happens to correspond to the set of measures which Eysenck found most powerful in distinguishing neurotics from non-neurotics. It loads most highly the tests of sway suggestibility stemming from Hull, the classical motor-rigidity or perseveration measures of Spearman, the test of attitude fluctuation which we proposed in 1943 [2], and also the usual measures of high level of aspiration, of preference for form over color, and of a high ratio of the psychogalvanic response to signs of threat as contrasted with reaction to the nocive physical stimulus itself.

If you feel baffled of immediate insights into the underlying principle governing the variables picked out by this factor I beg you not to feel worse than the woman in the well-known zoo story who saw her first giraffe and remarked, "I just don't believe it." One would be disappointed in the method if it turned up only patterns that we already know. The method of factor analysis seems to some experimenters peculiar in that it does not initially require that the experiment start with anything but the most general hypotheses. In fact it concurs with the position in Spence's paper, that it is better to begin looking for lawful relationships than for verifications of unnecessary hypotheses. Factor analysis first presents us with some dimly perceived pattern, in which initially none of the variables is very substantially loaded. But at that point we must form a slightly clearer hypothesis, by inference and abstraction, from the given variables. This is next tested by entering a new factorization with variables more pointedly chosen to represent that hypothesis, and so the cycles proceed [*1*, *6*, *17*].

Incidentally, if any personality factor turned out to load very highly in our first tests a small set of variables which *very closely resembled one another*, I should be suspicious that we had picked up some very specific behavior pattern—some too narrow group factor, in factor-analytic terms. A major personality factor should affect a very wide range of behavior, and the factor we have mentioned is satisfactory in this respect, for it covers operational responses which can be defined by such apparently diverse operational terms as instability of attitudes, poor will control, lack of realism, and a tendency to project emotional values into stimuli in an irrational fashion. These subsume into a single principle only if one takes a sufficiently penetrating theory of personality, but their pattern obviously resembles that *C* factor of ego strength versus emotionality which we have seen has been obtained also in clinical and behavior ratings.

Turning to a second example of these independent factors or source traits, we may take the loading pattern [5] comprising slow speed of perceptual closure, tendency to perceive threatening objects in unstructured pictures, large mean magnitude of general psychogalvanic response, absence of questionable reading preferences, high index of carefulness in following instructions, and much slowing up of reaction time when more complex choice reactions are demanded. Here we have hypothesized that the common principle is one of inhibitory tendency, which, when stronger, makes for greater caution in coming to perceptual closure, and so on through the other variables named. This has been tested to some extent by including in the succeeding factorial experiment a test of maze learning under reward and under punishment, in which the individuals high on this factor were found more responsive to punishment. The whole pattern suggests that this is the equivalent of the surgency-desurgency factor in ratings. But questions of identification with behavior rating factors, as well as of interpretations through new variables, cannot be considered settled until independent workers have systematically checked by further factorization. At present all that we can be sure of is the outline of certain factors, and we do have fairly substantial evidence that they match certain rating factors and have a certain character.

Before asking what light the discovery of these patterns may throw on the nature of the learning processes through which the organism passes, I must beg time to remind you of a new contribution from this direction to the determination of relevant structure, namely, the evidence about motivation structure [2]. The application of factor analysis to motivation and interest variables was initially a forlorn hope of structuring that which had hitherto defied objective methods as evidenced by the writings of Murray, McDougall, Freud, and Tolman. Although the need was great we had grave doubts whether the method which had produced a few stable patterns in other realms of personality would be able to deal with the complex, subtle, and changing relations of this motivation realm. But in fact it succeeded beyond our most sanguine hopes [3, 8, 9].

Fifty diverse attitude-interests had been measured on 200 young men by objective methods, that is to say, not by opinionaire methods or self-evaluation but by utilizing the magnitude of the PGR response, and by the application of perception and learning principles such as retroactive inhibition and attention to cues. The intercorrelations of strengths of motivation were factor analyzed and yielded patterns that were clearly those of primary drives [2, 3].

We had hypothesized that we should find both primary-drive patterns and sentiments, and had consequently put into the fifty interest-attitudes a pair of marker variables, i.e., operational representatives, for both drives and sentiments. I use "sentiments" here in the same general sense as in the accompanying paper by Adams, and like him I regard it as of central importance for learning. However, I make a slight distinction between sentiments and attitudes in a way which he does not [2]. I agree that a chain of attitudes can be recognized in which each is instrumental (or "subsidiated") to the next, but on closer observation it will be found that these interlock in a "dynamic lattice" [2]. What I call a sentiment about an object is a point in this lattice where several attitudes become integrated [2]. In regard to such structures the factor-analytic evidence was puzzling, for we obtained evidence only for the pattern of the self-sentiment, that is to say, of the integration of attitudes around the self-concept. It may be that the covariation of interests centering on such social institution patterns as religion, the home, and patriotism is relatively slight. Until the study has been repeated a third time, with special attention to magnifying this part of the picture, it would, however, be premature to deny the possibility of existence of such acquired patterns of attitude integration.

In order to escape that unearned increment of unjustified interpretations which every psychologist feels entitled to apply to the term "drives" we have used the term "ergs" for the empirically established patterns of hunger, sex, fear or anxiety, dominance, etc., emerging from the factor-analytic operations. This term avoids also the current semantic confusion of "drives" with "drive." For the present an erg means nothing more and nothing less than this immediate motivational pattern, and it remains for further research to determine, for example, whether the strengths of these nine or more ergs, when examined by twin studies or by physiological research, are largely innately determined or not. However, their similarity to patterns in primates, and in mammals generally, certainly suggests that we are dealing with inherited tendencies. Our finding of a very definite erg of curiosity or exploration certainly agrees, for example, with the convincing motivational, problem-solving studies of Harlow [*11*], who shows that monkeys, like professors, will put in a sixteen-hour day without such gross rewards as food and drink. Parenthetically these researches also support Harlow's questioning of two other pieces of semantic nonsense, namely, calling some drives, such as curiosity or fear, "secondary" to such "primary" drives as food-seeking or sex, and calling all drives "anxiety." I presume that when "instinct" or "drive" became socially unpopular someone looked in a dictionary and found "tension" was an acceptable synonym for "drive," and a little later some second unsophisticate looked in a dictionary and found "anxiety" as a bright new synonym for "tension." Anxiety, if words are to have any meaning, is surely properly a special derivative [2] of the fear or escape drive. But our results clearly show eight other drives beside escape, and so, perhaps, following Harlow's apt witticism, we shall have to speak of no less than eight "varieties of unanxious anxiety."

If these ergic structurings among objectively measured attitudes are confirmed by further, independent researches it will be seen that we have made two advances which should be of especial interest for learning experiments. First, we have arrived at a pattern of weighted elements by which the strength of a given drive or erg can be estimated in a human being or a lower animal at a given moment in a given stimulus situation, and second, we have made possible the alternatively oriented analysis of stating quantitatively what particular final goals or incentives are concerned in motivating any given attitude or habit that we encounter in action in nature.

The latter method of analysis has been expressed in what I have called the ergic principle of attitude measurement. Initially an attitude may be defined as an interest, within a given stimulus situation, in following a certain defined course of action. One may put any attitude into the paradigm, "In these circumstances I want so much to do this with that," which defines, as I have indicated elsewhere, four conditions: namely, the stimulus; the direction of action taken by the response; the object, if any, manipulated by the response; and the strength or urgency of the response itself. This basic definition of attitude, incidentally, is one which social psychologists have difficulty in grasping, a fact which seems to me to have some correlation with their polling prediction difficulties, but learning theorists, I believe, will find it rooted firmly in familiar concepts.

Now the ergic principle of attitude measurement abstracts from the measurement data according to a certain mathematical model primarily designed to permit a number of useful further calculations to be carried out. It expresses the attitude as a vector quantity, the length of which represents the *strength* of the interest in the course of action and the direction of which, defined angularly in relation to the known ergic coordinates, expresses the *quality* of the interest, or, in other words, the ergic goals satisfied by that course of action. Where normative units of measurement can be employed, as in *R*-technique studies, this formulation permits us to add the attitudes of different people, as, for example, in determining the total attitude of members to a group, i.e., its synergy [2] in group dynamics. Where ipsative (*and* normative) units are employed it permits us to work out measures of conflict within the individual, of the strength of ergs excited in a given stimulus situation, or of the relative investment of interest in different learning tasks.

I spoke a moment ago of the uncertainty we had felt regarding the capacity of factor-analytic methods to structure such complex and fragile data as motivation measurements, in which methods of objective measurement are still tentative and where transient and oscillatory conditions of physiological state and stimulus situation intrude into the variance as much as do the fixed characteristics of the situation and organism, reducing reliability coefficients on any variable measured. But here the method of *P*-technique factor analysis, which we had developed with clinical purposes in mind, actually made capital out of the very disadvantages of the situation. When we turned it upon the problem, it brought out in still clearer terms the ergic motivation structure just described as resulting from the *R*-technique studies.

P-technique measures the same population of attitude-interests on one person, day after day, for, say, a hundred days. The changes in stimulation can either be experimentally arranged or left to the impact of daily events. Although correlation of the variable series and subsequent factorization yielded, as stated [9], the same ergs as had been obtained by *R*-technique studies and also the same self-sentiment structure [8], the exact relationships were now naturally those of the unique trait rather than common trait patterns. The personal conflicts in the individual, as revealed by the positive and negative loading antitheses, when compared with those inferred clinically, showed excellent agreement, both as to magnitude and as to the ergs and personality structures involved. Further, a calculation, from the ergic loading patterns, of the strengths with which ergs were stimulated, when plotted over the hundred days, showed decided changes in level very pertinently related to the events in the young man's life. For example, the measured fear or anxiety drive climbed rapidly around examination time, and the erg of tender protectiveness (perhaps what a rat psychologist would call "maternal drive") rose strongly during the week that the subject's father was seriously ill in the hospital. Incidentally anyone particularly interested in method will have recognized that *P*-technique, especially if used in what I have called the *condition-response design* [6], may deal with causal sequences of stimulus-response and thus transcend the preoccupation of *R*-technique factorization with cross-sectional relationships only. This will be important in our later discussion of methods capable of combining the perception of patterns provided by *classical factor analysis* with the perception of stimulus-response relations provided by *classical univariate experiment,* as in learning theory studies.

Now I have taken some of the time up to this midpoint in my talk in a detour aimed briefly to describe the foundation of systematic conceptions and the methodologies of measurement in personality and clinical psychology which in my opinion can be most profitably related to learning theory. Such a review has been indispensable because much of the work is too recent to be widely circulated among specialists in other fields, and yet, without awareness of its outlines, it would be impossible to proceed to attempts at exact integration with learning theory. As I have indicated above, one cannot be very enthusiastic about that variety of learning theory which develops in a vacuum as far as personality is concerned and remains vacuous of anything but logical constructs invented in an armchair. Would it not be better frankly to recognize this as philosophy?

Now, as suggested a moment ago, the obvious and immediate help which personality study can give to learning experiments lies, first, in pointing out the structures which learning needs to explain, and second, in providing sound measures by which the increments in these structures can be determined. This applies both to learning and to that relearning which the clinician calls psychotherapy. In regard to this latter it should be noted that the means are now available, through basic research in personality, of objectively measuring personality during changes in psychotherapy. Preliminary results indicate, as one might suspect on theoretical grounds, that change occurs not just along one supposed highway of "getting cured" but simultaneously with respect to several dimensions of personality. For example, in therapy, change occurs in the direction of increased surgency (factor *F*), increased integration (factor *G*), and decreased free anxiety (factor *O*). It is interesting that the changes which Petrie [*16*] has measured in lobotomy follow this same course, with, in addition, a decline in factor *B* or general mental capacity. Somehow the therapeutic relearning process produces increased integration and decrease of general inhibition other than by the severing of neural connections involved in lobotomy.

This example, dealing as it does with very massive reaction patterns in the personality—namely, those concerned with the degree of integration, the general level of inhibition, and the general proneness to anxiety responses—brings out very clearly the difficulties we now have to face in utilizing in personality experiment that learning theory which has as yet been untutored by any substantial contact with personality theory. Both the personality theorist and the clinician are bound to object, if they are frank, that the preoccupation of the learning experimenter with "rats and reflexes" has meant altogether insufficient regard in learning theory either for the personality patterns which are of prime concern or for the psychometric problems of accurately measuring personality change along such dimensions. For, in the first place, the personality factor patterns which make civilized man what he is, resulting from his exposure to institutional patterns, have rarely or never been investigated or demonstrated in laboratory animals. Second, as regards any attempt to measure such patterns, learning experiments seem to have been designed on the assumption that no attention need be paid to changes in the animal other than those in his time of running the maze or acquiring a reflex or whatever other *single* variable has been measured.

If we are really to get personality and learning theory together, according to the excellent intention of this meeting, I think the above-implied divergences in viewpoint and method upon which we have so far been stumbling only as uncertain innuendoes and embarrassing discrepancies should now be explicitly repeated and summarized, and this I shall do also under five main headings. As a personality theorist, who has done whatever he is capable of perceiving to rectify gaps in personality theory, I shall definitely adopt the position, for the sake of clear argument, that the remaining perceptible shortcomings reside in learning theory. In so doing I am not quarreling with the position taken in Spence's paper, namely, that learning theory inevitably began with the simplest manifestations and that for the sake of morale it was good to get some lawful relations as soon as possible at this level. I am only stating as a present fact that learning theory has not yet grown *beyond* that point, and cannot therefore integrate as well as one would like with personality theory. But I shall not be surprised if the learning theorist, properly aware that his work is initially well founded, replies, like the Kentucky horse dealer accused of selling a beast with defectively short legs, "I see nothing wrong with them. They reach down to the ground." But I am merely asking whether they can lift the body of personality theory to its proper level.

1. First, my criticism that learning theory has concerned itself with individual response variables, instead of factor patterns and their measurement, may be resolved in part into a charge of neglect of *the organism as such.* Early learning theory simply considered response as a function of stimulus, perhaps with perceptual constructs in between, thus:

$$R = f(S) \tag{1}$$

In response to the raised eyebrows of personality theorists some, but not all, learning experimenters have formulated response as a function of stimulus *and* organism, thus:

$$R = f(O, S) \tag{2}$$

but the rigidity of thought is such that this has never been done wholeheartedly; and the organism, for many learning theorists, is still largely a vacuum, or a disembodied "state," or a rather vague conceptual area populated by intermediate constructs having to do with *antecedent external stimuli.* These concepts are incapable of dealing with the full facts of individual differences, and they also miss some indispensable general characteristics that the general organism should bring into the equation.

Regarding the first, the learning theorist often takes up a very clear position, but it is the position of Wundt in 1880 when he told McKeen Cattell that he was interested in *common processes,* not in individual differences. Surely we have learnt by now, however, that the best way to find out the nature of common processes is to watch them varying, just as the way to perceive a rabbit in a field is to wait till it moves. Surely it is understood among us that only such amateurs as novelists, applied psychologists, and parents are interested in individual differences for their own sake, and that our scientific aim is to use them only as the best avenue to general processes and truths. Consequently I repeat that one of the methodological advances in learning experiments that could result from the present contact with personality theory would be the effective utilization of individual differences and multivariate methods to identify and measure the intermediate variables between stimulus and response in the above equation. In the simpler sense this step would merely introduce into the equation the organismic variables, in which the learning theorists are to some extent entitled to say they are not particularly interested —variables such as constitutional levels of intelligence, rigidity, and drive strength. (Though I observe that such learning researchers as Maier [*14*] do not share in the general neglect of these factors.) But even granting their right only to be interested in the non-organismic, non-particular factors, the learning theorists surely need to know how the organismic factors interact with the non-organismic and what the relative contribution of each is to the total variance in the learning process.

However, what I want to stress is that there is a second application of this individual-difference approach which could yield the hypothetical constructs in which learning theory is *primarily* interested, namely, the use of individual differences in previous exposure to stimuli, in time intervals between learning experiences, in delay of reward, etc. It may be replied with astonishment that these are the very substance of the learning theorist's daily bread. But this reply misses the point, for two reasons: first, he at present introduces these condition variables one or two at a time instead of in a multivariate factor-analytic design, such as could yield good definition of factors; and second, he proceeds, in analysis of variance and similar designs, as if he were quite unconscious that such variables may be correlated. I should like to see a closely reasoned discussion of a multivariate design to investigate factor analytically the reality of the hypothetical constructs in the Hullian equations, such as habit strength, inhibitory potential, and excitatory potential [*12*], for the present univariate designs seem to me to be working with too many unknowns in the experiment to permit the net of deduction to close in effectively on so many constructs from so few variables.

You will notice, incidentally, that I am indicating theories in terms of experimental designs rather than in terms of the familiar theory labels, but I do so deliberately because I think we perhaps talk a little too much in terms of half-understood generalities and suffer from the reification of names that may have no operational justification. Here I strongly endorse Adams's remark that we seem to have passed in one generation from the barbarism of brute empiricism to the decadence of an excessive concern with self-conscious theorizing.

Now I should like finally to put before you the personality theorist's development of the $R = f(O, S)$ equation and suggest to you that the most fertile growing point for theory today would be a side-by-side discussion of this formulation on the one hand and the Hullian or the Spencean formulation [*18*] on the other. If you will permit me a metaphor of spring, I am hoping that the seed of the specification equation will become pollinated by a few constants wafted from the learning equations. The personality specification equation [2] states that a learning performance, R, in a stimulus situation, j, can be written thus:

$$R_j = S_{1j}E_1 + S_{2j}T_2 + S_{3j}E_3 + S_{4j}T_4 \cdots + S_{nj}T_n \qquad (3)$$

where the S's are situational indices, obtained as factor loadings, and the E's and T's are, respectively, the strengths of motivational and non-motivational traits peculiar to the organism and its previous learning experience. It will be observed that the S's are the dimensions of the stimulus situation, giving substance to S in the basic equation (2) above, while the E's and T's give the dimensions of the organism, including its past experience, and thus permit a realistic representation of O in the basic equation.

Time shortage, if not more serious shortages, stops me on the threshold of discussion of the integrational possibilities. I might point out, however, the resemblance in meaning between any one of the product terms and the excitatory potential of Hull's scheme, considered as a similar product of drive-stimulation strength and habit strength. T may be considered the strength of the habit and S the capacity of the situation to stimulate drive. Since the personality theorist cannot assume that any animal, least of all a human being, is acting from one pure motive, the product terms differ from the

Hullian formulation here in that they repeat themselves as shown, for each of the various established drives in the organism.

Although I have charged that the first difficulty in getting together with learning theorists is that they have left out the organism, which to my mind is as bad as producing the play without Hamlet, perhaps it is not too late to adopt a Dale Carnegie gesture and put this criticism in terms of a much more minor peccadillo. Indeed, if in this matter I confine myself to what learning theorists actually *do*, or rather fail to do, instead of to what they *think* (or what I think they think), then the objection would reduce simply to what I have stated before—namely, that they represent the organism, and everything else under heaven, by a single response instead of by those patterns which the general psychologist, the clinician, and the sociologist regard as important. What is worse, they try to infer its nature from this single response. I suggest that, quite apart from the interest and importance of the total patterns, we cannot assume that the laws of learning themselves, as they apply to such patterns representative of the total organism, are identical with those found for single variables.

Perhaps it should be made quite clear at this point that because I emphasize patterns among responses there is no reason to claim, as some do, that I must therefore neglect the relation of responses to stimuli. In a well-known paper some years ago Spence criticized [*18*] the methodological defects of those theories and mental measurement approaches which seem interested only *in relating response to response,* whereas the relating of response to stimulus alone, he pointed out, gives us those causal sequences in which science is ultimately interested. His observation is a penetrating one, and I certainly would not stop to defend those item-analyzing educational psychologists and psychometrists whose interest finishes at correlation and who show no further interest either in learning theory or in general personality theory. But I would at least say that as far as I am concerned the cross-sectional patterns are only preliminaries—though *very essential preliminaries*—to sequential study.* For good cross-sections are as essential to intelligently planned human learning study as are well-focused individual pictures to clear perception of action in a movie film. However, we do not always *have* to separate cross-sectional from longitudinal interests, for there now exist three methodological contributions from personality research which inaugurate

* In these, essentially, we are holding the stimulus constant and seeing how varying reactivities produce the pattern—to conceive this in classical experimental terms.

simultaneous study of cross-sectional patterns and causal sequences—namely, (1) *P-technique* [*4*], (2) the *factorization of increments* [*6*], and (3) that combination of factor analysis with factorial methods which I have called *condition-response factorization* [*6*].

2. My second main suggestion for integration concerns the concept of ergs and the methodology of measuring drive strengths. For I trust that the introductory illustrations suffice to support the argument that we need to shift from single variables to patterns not only in terms of response measurement but also in terms of measuring the manifestations of the independent variable—the drive strength. It is inadequate and unnecessary to restrict the measurement of thirst to hours of water deprivation or of fear to the measured voltage of a shock. The work I cited shows that with human beings the number and the nature of independent ergs can be determined, and that one can objectively discover the weights by which the strength of excitation of a drive can be estimated from a measured set of manifestations. Anderson's pioneer work on factoring the reactivities of the rat points to the practicality of similar substitutions of pattern for single variables there also.

If we are to be scientifically cautious we must admit that the learning curves and laws obtained for differences of ergic strength, as thus more completely conceived and measured, *could* be significantly different from those so far obtained on a univariate basis. Even if they do not turn out to be different in general form, as some of Harlow's observations on animals suggest, they will almost certainly be more accurate. Furthermore, such an approach could open up the whole question of whether the differences of *quality* between drives are associated with important differences in the process of learning and forgetting. For example, it is implicit in Freud that neurosis and repression occur only with appetitive, viscerogenic drives; and Brozek's demonstration of human experimental neurosis through hunger still leaves positive evidence only for the efficacy of this same drive category. Is it possible that forgetting is a different *kind* of process with non-viscerogenic drives such as fear, curiosity, and gregariousness? With objective delineation of the drive structures and the determination of weighted composites by which they may be measured, research on this question becomes possible.

3. Next, in the discussion of *rapprochement,* I want simply to ask a question rather than make a suggestion, for here the learning field seems too semantically confused to make any intelligible suggestion without much time spent on defining terms. The question is this,

"How can we get together when personality theorists, following clinicians, conceive learning as a dynamic process, in terms of goals and rewards, whereas for a substantial number of learning theorists the paradigm of learning always seems to be primarily the conditioned reflex?" I suppose the answer is that we can get together only with learning theorists like Hull, Thorndike, and Spence, who bring conditioning under reward learning, or with two-factorists like Skinner, Mowrer [*15*], and Maier. But if the position of the latter should prove to be scientifically sound then the personality theorists and clinicians are grievously neglecting a principle that must have more representation in their data than they have yet recognized—and personality theory will have to be reformed accordingly.

4. My fourth suggestion regarding our integration difficulties is that we should be prepared to admit in our diagnosis that the gaps are sometimes wide because they are historically old and deep. Let us be clinicians and try to remove the dissociations by facing our past traumata. Historically, in the nineteenth century, the pukka sahibs of scientific psychology were unquestionably to be found in the laboratory, rather than in the clinic, and within this shrine their fetish was the brass instrument and the univariate, controlled experiment—sacred relics from physics. To this day many psychologists in perception and learning seem not to have realized that this particular ritual is not the whole of scientific method but only one narrow branch of it.

Now I have no intrinsic objection to this scientific "holier-than-thou" attitude so prevalent among learning theorists—if it would not interfere with their learning. Parenthetically I have no doubt that they are indeed more scientifically respectable than, say, psychoanalysts, and their attitude is better than the "more-popular-than-thou" attitude of the clinicians or the "richer-than-thou" attitude of industrial psychologists. But what these scientific pharisees or hermits (according to your view) fail to realize is that in many investigations the new multivariate methods, which psychology itself may justifiably feel proud to have given to the sciences, have better claims to scientific penetration and rigor than the imitations of physics with which the old guard has so rigidly been practicing.

To all but these Rip Van Winkles it has been evident for a generation that important human learning situations, such as those with which clinical and social psychology have to deal, cannot be dragged intact and alive into the laboratory. The controlled, univariate experiment, in which nothing but the independent variable alters to an important degree, becomes inapplicable and obsolete—if only because of the limits to our rights of major interference in human lives. If at that point you can think of nothing better and you prefer the shadow of scientific method to the substance of living psychological events, you can continue the univariate experimental procedure with animals; but this, as I think Mowrer has found, is a cul-de-sac which is soon exhausted, for you cannot reason very reliably by analogy from rats about human social problems. An alternative scientific method had therefore to be conceived, in which we let events happen in life as they will and tease out by statistical finesse what cannot be handled by brute experimental control. We are indebted to such men as Spearman [*17*], Fisher, Thurstone [*19*], and Burt [*1*] for most clearly perceiving this need and starting to hack a new path away from the beaten but misleading track of the classical experiment.

But multivariate statistical designs do more than provide an effective way of handling what used to be called the controlled variable and the uncontrolled variable in situations where control is impossible. The factor-analytic approach [*6*], for example, in addition tells us first where much of the unknown error variance is located, and second, much more about the operations by which our concepts are defined.

We must look at this last point more sharply, for it contains the crux of an issue which is deeper than that of experimental design alone and which has been hovering in the background of my earlier points. May I assume that we are all good operationists together and that the established meaning of a concept goes only as far as the operations invoked to represent or to test it? Now the univariate experiment leaves the meaning of a concept hitched to a single operation. For example, hunger is hours of deprivation of food and emotionality in a rat is the number of defecations in an open space. I submit that when the sharp winds of argument about "meaning" begin to blow it is much better to have a concept tied down at the corners by several variables than by some isolated operation. When a personality dimension or motivation strength is hypothesized as a whole pattern of operations of measurement, not only is it more effectively defined and more conclusively tested by experiment but it also provides a firmer basis for an architectonic building of further research, as a table with several legs is a firmer foundation than a one-legged table.

Or, to look at this point another way, there are an infinite number of individual variables by which one might choose to represent some concept such as rigidity or ego strength or anxiety or drive strength. History unfortunately shows that no two successive studies will use the same variable for the same concept, for the modern psychologist might lose his individuality if not his soul by such obvious lack of originality. Indeed the capacity of a finite number of psychologists to spread themselves over an infinite number of variables is something for sampling theorists and even mathematicians to marvel over. But the well-known upshot of this marvel is that no two studies integrate, and that where chemists make architectonic progress, by deigning to recognize the same elements, psychological research is all too frequently circular or inconclusive. If psychologists would deal with established patterns among variables in a given area, as their reference concepts in research, this difficulty would vanish.

An illustration of what happens in this respect when univariate learning theorists get loose in the multivariate realm of personality is provided by half a dozen attempts to do laboratory experiments on the relation of such matters as the self-concept, or personality levels in anxiety, to such laboratory measurements as rate of extinction of reflexes, rate of closure, or speed of reaction time. The contrast between the impeccable scientific design developed in the laboratory measures and the uninformed plan for the measurement of anxiety is a painful one. If we are to build bridges from personality to learning theory why not build them expertly at both ends? In these cases the learning theorists have founded the bridge on an admirable pier of solid experiment at their end, but have constructed the other pier—that concerned with defining and measuring anxiety—out of an amateurish heap of rubble. They have had no better conception than to throw together sets of questionnaire items having "face-validity," which, as we all know, is a polite term for no validity at all. I submit that an examination of the evidence on personality factors would have indicated the following three points: (1) that though available data show only one erg of fear or escape [3, 9] there are indications that its manifestations become separated into four distinct sources [2] in personality structure, namely, those covered by the factors labeled *F*, *M*, *O*, and *Q4*. As far as interpretation is now possible these correspond to an anxiety level of general previous punishment, an anxiety of superego action, a free-floating "anxiety hysteria" anxiety, and a somatic anxiety like that defined in anxiety neurosis. Almost certainly the associations of the composite with conditioning are due to only one of these, and the researches in question will need to be done over again to show which one, and to make an intelligible interpretation possible. Incidentally, the extraordinary semantic carelessness which says anxiety level is only drive level would still further confuse the design and interpretation of such experiments: the fact is that the questionnaires used have dealt with anxiety, not drive; (2) that furthermore the above four independent forms of anxiety in personality are in some cases measurable by more objective devices than the questionnaire; and (3) that some indication could have been obtained from previous work, notably the observations of Stephenson on the resemblance of extinction phenomena to the classical motor-rigidity factor measures, and the extensively investigated relations of this rigidity factor to personality, as to which personality factors, anxious or non-anxious, are more likely to show significant relations to conditioning extinction.

5. I think that two of the four just causes and impediments to the wedding of learning and personality theory that I have pointed out may be considered to be theoretically complex and debatable, and the other two obvious and remediable. My fifth and last impediment is, I hope, simply remediable though theoretically complex. This particular obstacle to bringing personality theory and learning theory together resides in the fact that in human learning the rewards are often so obscure, so rooted in unconscious rewards (such as those to superego value systems), or so hidden in faint cues in the social environment, that any attempt to relate drive strength to learning in such situations, in the honest fashion of rat and laboratory experiments, appears to be denied a chance. To add to the complication the proportions of the standard "varieties of learning" in these human fields of experience are in general very different from those to which we are accustomed in the laboratory. On the question of varieties of learning [2] I presume we can agree descriptively on three: (1) conditioning by contiguity, (2) rewarded means-end learning, and (3) rewarded "integration" learning—though some would run the second and third together. However, though they are both forms of dynamic learning, the first of them concerns a single drive and only changes in the means to its goal attainment are involved, whereas in the second there is the larger problem of conflict resolution. This integration learning, moreover, is the principal concern of the clinical psychologist. Briefly it may be defined by the fact that conflict ends in the denial of some individual drive goals in the interests of a greater reward to the total needs of the organism.

Now the specification equation analysis of any given learning performance into situational indices on the one hand and ergic and metanergic personality structures on the other—as stated earlier—

actually provides us both with the information we need concerning the sign-significate satisfactions existing in the obscure total stimulus situation and also with a statement of whether the learning is largely a means-end learning or an integration learning. The representation of the first requires no amplification: it has already been seen that the S's define the motivational dimensions of the stimulus situation. Regarding the nature of the learning I shall simply point out for further reflection that the situation involves integration learning to the extent that there are conflicting positive and negative loadings to the situational indices, particularly in respect to the self-sentiment and primary drives. A difference of sign means that a positive satisfaction of one erg is to be obtained only at the cost of suppression of another. Some interesting examples have been discussed elsewhere [8, 9].

Regarding the revelation of the cues that lie in the situation, it has just been stated that magnitudes of the situational indices attached to particular drives indicate (with certain statistical modifications) the extent of the provocations to those drives somehow hidden in the situation. The next step in proceeding from this basis would be to locate and manipulate the specific stimuli or aspects of the situation concerned, whereupon we might hope to obtain, by the ensuing learning experiment, an experimental check on the factor-analysis S values. A methodological alternative would be to *accept* the strengths of excitation given by the factorization and attempt to establish laws of learning by relating the emergent personality patterns to these accepted ergic strengths.

In conclusion it is my impression that I have rarely seen different specialist fields brought together in which the hopes of substantial gains from interactions are so great. There exist at any rate two substantial and precise predictive systems: first, the crystallization of facts and laws in factor-analytic measures of personality, and second, the equally effective crystallizations in learning theory. Since both are firmly rooted in experiment there must be some way of linking one with the other in a further set of lawful relationships. For researchers who aspire to make those links there now exist some potent suggestions in methods and hypotheses, which can be utilized by those bold enough to think in new ways and patient enough to study the theoretical implications of what has been offered. One reason why so many possibilities have been unearthed here is that this meeting of specialists has been long overdue, and the University of Kentucky is to be congratulated on the initiative of its psychology department in at length bringing this synthesis about.

REFERENCES

1. Burt, C., *Factors of the mind,* London, University of London Press, 1940.
2. Cattell, R. B., *Personality: a systematic, theoretical, and factual study,* New York, McGraw-Hill, 1950.
3. Cattell, R. B., The discovery of ergic structure in man in terms of common attitudes, *J. abnorm. soc. Psychol.,* 1950, *45,* 598–618.
4. Cattell, R. B., *P*-technique, a new method for analyzing the structure of personal motivation, *Trans. N. Y. Acad. Sci.,* 1951, *14,* 29–34.
5. Cattell, R. B., A factorization of tests of personality source traits, *Brit. J. Psychol.,* Stat. Sect., 1951, *4,* 165–178.
6. Cattell, R. B., *Factor analysis for psychologists,* New York, Harper and Brothers, 1952.
7. Cattell, R. B., Research designing in psychological genetics with special reference to the multiple variance method, *J. human Genetics,* 1953, *5,* 1–21.
8. Cattell, R. B., and Miller, A., A confirmation of the ergic and self-sentiment patterns among dynamic traits as determined by *R*-technique, *Brit. J. Psychol.,* 1952, *43,* 280–294.
9. Cattell, R. B., and Cross, K. P., Comparison of the ergic and self-sentiment structures found in dynamic traits by *R*- and *P*-techniques, *J. Pers.,* 1952, *21,* 250–272.
10. Eysenck, H. J., *The scientific study of personality,* New York, Harcourt, Brace, 1952.
11. Harlow, H., see his article in this volume.
12. Hull, C. L., Mind, mechanism, and adaptive behavior, *Psychol. Rev.,* 1937, *44,* 1–32.
13. Hull, C. L., *Principles of behavior,* New York, Appleton-Century, 1943.
14. Maier, N. F., *Frustration: the study of behavior without a goal,* New York, McGraw-Hill, 1949.
15. Mowrer, O. H., *Learning theory and personality dynamics,* New York, Ronald Press, 1950.
16. Petrie, A., *Personality and the frontal lobes,* London, Routledge and Kegan Paul, 1952.
17. Spearman, C., *Psychology down the ages,* London, Macmillan, 1937.
18. Spence, K. W., The nature of theory construction in contemporary psychology, *Psychol. Rev.,* 1944, *57,* 47–68.
19. Thurstone, L. L., *Multiple factor analysis,* Chicago, University of Chicago Press.

Prospects and Perspectives in Psychotherapeutic Theory and Research

JOHN M. BUTLER

It is quite apparent to anyone who reads the literature on psychotherapy that the level of discourse is low and that the different schools of psychotherapy have different systems of propositions. These systematic differences make intercommunication with regard to single propositions extremely difficult and in so far as science involves good intercommunication demonstrates the low level of scientific understanding in this domain. It will be my purpose in this discussion to consider how, in a domain in which all seem to profess a scientific orientation, this low level of communication came into being and to consider how the domain of learning can serve to provide a common frame of reference which would promote communication as between systems of psychotherapy.

In reviewing the modern history of psychotherapy, which as far as I am concerned dates from the early days of Freud and Janet, he who has the advantage of hindsight and contemporary scientific sophistication sees a fascinating spectacle. He sees a host of brilliant men inspired by the success of physical and biological science, bending their energies to the creative formulations of the processes of psychotherapy and of the nature of human nature. These men, who worked in a restricted and really rather simple social situation, that of psychotherapy, which had the unique feature of being safe for the therapist in a way in which everyday life was not, were therefore in a position to make observations which were relatively objective, i.e., were not ego-involved. And they were, it should be noted, with a given individual for a considerable span of time. Today it is clear enough that, although these circumstances permitted relatively objective observation, still the behavior of these therapists was in itself influential in producing the phenomena (and their time sequences) from which theories of psychotherapy and, inevitably, theories of human nature arose. It is also clear enough that no two of these men were identical or exhibited the same behavior pattern in therapy.

Consequently it seems quite evident, to one who is looking back, that different observations would be made, and that different formulations of psychotherapy and personality theories would emerge. It is also to be expected that these men would make observations and evolve theories which would be quite unacceptable and upsetting to their contemporaries. And indeed it is true that even now psychotherapists are often subject to severe attacks because of the threatening nature of their conclusions. Seen in perspective then, it is quite natural that these men who believed in the accuracy of their observations, who were convinced of the importance of their work, and who were the objects of severe attack, would be intolerant of opposition both from the public and from their colleagues who, on the basis of different observations, differed theoretically.

The basic point here is that in their time no basis for settling the to-be-expected differences existed, even in principle. These men were endeavoring to take a scientific attitude in a domain in which very few, if any, scientific studies were possible. What were the techniques applicable to the scientific study of human behavior in 1900, for example? And, for that matter, where existed the concept of the study of behavior in 1900? What theory of science existed in 1900 which would help the therapist to even formulate a scientific investigation of his problems? It is certain that Pearson's "grammar of science" could not aid the therapist; neither could the then current psychological theories (though they could and did contribute to therapeutic technique) nor the discipline of statistics.

We expect then, under such circumstances, to find these able and gifted men differing deeply in their approaches and theories, convinced of the general importance of their work with little in the way of scientific methods to aid them in settling their differences. And we also expect them, in such circumstances, to settle their differences in the ancient ways: by dismissal, by undermining, by expulsion from the group, by leaving the group, and, above all, by impassioned persuasion. It is clear that for the most part such social processes determined the acceptance and rejection of therapies at least as much as their validity. Who is to say whether "individual psychology" is better or worse or more valid than psychoanalysis, on the basis of scientific evidence? Yet it is fairly clear today that Freud left more intellectual heirs than Adler, and that as a group enterprise psychoanalysis is more successful than "individual psychology" in terms of number of publications, publicity, and adherents, both practicing and devotional. As for those who did not join groups, did not have the temperament for scholastic disputes, who remembers them?

As we look to the theories evolved we see that some are more consistent than others, some *seem* to make more sense than others, and some have broader scope than others. But as we look at the evidence we see what are called in the area of student personnel work "systematic anecdotal records." We see accounts of the behavior of patients fitted into pre-existing frameworks with certain hours picked out for analysis, with an implicit rather than explicit coding of behavior; we see perhaps hundreds of therapeutic hours condensed into a few pages of description, followed quite often by really impressive theorizing which has no ascertainably direct connection with the basic data, i.e., with the coded behavior of the patient or patients. And this condition has persisted, with but very few exceptions, to the present.

With this perspective of the past it is hardly surprising to find certain kinds of vested interests developing over a rather considerable time period with as some of their consequences a general lack of interest in evidence, strongly grounded beliefs in the truth and validity of one kind of theoretical approach, and a general circularity of theory which is most discouraging to those who really want to approach the domain of psychotherapy with both a scientific attitude and a scientific method of implementing that attitude.

To be quite frank, it is my opinion that, dark as the perspective is, the present is quite as dark when one views the different extant schools of psychotherapy. What once represented an attempt at scientific formulations, or at least some kind of prologue to scientific formulations, now represent dogmas, now are limiting frames of reference, the boundaries of which it is dangerous to approach if one values his standing in his group and is unwilling to risk or to undertake the formation of a new group. The darkness of this picture of parochialism and intellectual provincialism is aggravated when one considers that today there are impressive general theories of science, highly sophisticated methods of statistical analysis which are consistent with scientific method, and a growth in psychological theory and experimentation which is largely ignored by psychotherapists who still grind out books on human nature, claim one theory is "better" than another, and claim priority for the validity of their theories of personality—all in the name of science. For myself, I see no way out of this situation from within. It seems to me that the dilemma must be solved from the outside, that the parochialism can only be broken down by reducing theories of therapy and therapy-based personality theories to more general psychological theories which will "contain" all of them, so to speak. The apparent contradictions between the propositions put forth by the different therapeutic schools can only thus, it seems to me, be really examined. And it seems to me to be almost certain that the therapists are not going to do this themselves. At least they show few signs of wanting to.

I am of the opinion, at the moment at least, that this aim can best be realized through learning theory. Admittedly, learning theory is parochial in the sense that there are several theories of learning. There is a decided difference, however, between parochialism, if this is indeed the proper word to use, in the domain of learning and parochialism in the domain of psychotherapy. Students of learning have committed themselves to a common enterprise, the scientific enterprise. Therefore, learning theorists are compelled to listen to each other. Tolman cannot dismiss Spence's theories as superficial even if he wanted to. He must and does address himself to experiments supporting or refuting Spence's claims, and Spence must do likewise. And over the years both behavior theory and expectancy theory have accommodated themselves each to the other on the basis of experimental results. What we have seen develop is both a certain convergence and an increasingly clear discussion of crucial theoretical differences.

With therapeutic theory the situation is far different. So far none of the major therapeutic theorists appears to have felt compelled to notice his major rivals except to dismiss them as superficial and obviously incorrect when indeed they have not been ignored. Their parochialism rests on dismissal, rejection, and the assumption of biased and superficial observation.

I am not, therefore, disturbed when a behavior theorist or an expectancy theorist or some other learning theorist decides to translate propositions in the domain of psychotherapy into terms of the particular learning theory of his choice. The propositions of learning theory are, by and large, grounded in a theoretic-experimental system. I am not even disturbed if the theorist extends his theory in a consistent way, without experimentation, in order to accomplish this end. I am disturbed when he decides to restrict his consideration of the psychotherapeutic domain to client-centered therapy or "individual psychology" or "analytic psychology" or "conditioned reflex therapy" or "dianetics" or what not, because this or that therapy has "unquestionably contributed the most" or "I am most familiar with it" or "It is easiest to translate."

In my opinion such restrictions perpetuate the existing parochialisms in the domain of psychotherapy. Either conversion takes place because coordination or translation is confused with reduction, or an impressive step forward is rejected because propositions favored by one's own variety of therapy are apparently invalidated by the translation. Therapists are not prone to give up formulations based on personal, intimate, and hard-won experience.

Although obvious enough, this matter of translation should perhaps be discussed explicitly. When a learning theorist is able to find a certain consistency between the propositions of a given variety of psychotherapy and his own variety of learning theory, he obviously regards his learning theory as being better in the sense of science than the therapeutic theory he is translating. Otherwise the translation would be in the other direction. I assume then that what happens is that the relatively inexact propositions of one domain, psychotherapy, are being coordinated with the relatively exact propositions of another domain, learning theory. Now I would not deny that such translations are potentially worth while; in fact, I assert that they are. There are, however, two considerations to be kept in mind:

1. The translations contribute nothing to our knowledge of psychotherapy. The specifiable referents of the terms of learning theory reside in the experimental situations created by the learning experimenters.

2. The behavior of patients and clients and therapists in psychotherapy is as unspecified as ever it was, and this is very unspecified indeed. What are exact terms in learning theory then become as inexact as their semantic equivalents in psychotherapeutic and personality theory, and the connections between the propositions are connections which can be demonstrated to hold for one domain but cannot be demonstrated to hold for the other. To put it bluntly, the unification of the domains has by and large been on the semantic level. Once one has succeeded in finding a certain isomorphism between the propositions of therapeutic theory, the propositions of therapy-based personality theories, and the propositions of learning theories, with some paring down of the former propositions and some relaxation of the latter, one finds himself turning to experiments with animals and finding analogues. These analogues, I might add, have not contributed much, if anything, to learning theory. Why should

they when the experimenter started with learning theory, found isomorphisms, and came back to the domain of learning? They do not contribute directly to psychotherapeutic theories because the referents of the terms used in such theories are unknown in any scientific sense; if any direct contribution is to be made it is to be made to learning theory.

Perhaps I can illustrate these points by citations from current literature. Dollard and Miller [2] have written a most stimulating and provocative book entitled *Personality and psychotherapy*, by which is meant apparently behavior theory and psychoanalysis. On p. 257, unsurprisingly, we find that "therapy [psychoanalysis] takes time." Why? Because (1) anxiety is produced by free association; (2) facing anxiety is painful and exhausting; (3) a self-chosen pace (in psychoanalysis) is a slow pace, and as a result extinction of anxiety attached to verbal responses is a slow affair; (4) the therapist can best advance therapy by insisting on free association.

I take it that these statements can be traced back to the propositions of their approach to behavior theory if "anxiety" is accepted as a strong stimulus. Thus:

If anxiety is a strong stimulus,
and
if anxiety is produced by free association,
then any response or classes of responses which reduce anxiety will be reinforced. Thus free associating will tend to cease.

If the therapist insists on free association,
and
if the free associating arouses anxiety,
then the therapist becomes a set of cues which arouses anxiety.

I find this *if-then* sequence which seems to be implicit in Dollard and Miller's account to be entirely reasonable. I simply want to emphasize, which they do not, that the therapist in his behavior both reduces and arouses anxiety and that thus, by his own behavior, regardless of transference, arouses directly competing response tendencies. And this may go on simultaneously; the therapist may be reassuring, etc., in his manner while insisting on free association. It follows from this that extinction of extinction is going on in their description of psychoanalysis with the balance being on the side of anxiety reduction in successful cases.

Now take a proposition apparently contradictory to the one that "therapy necessarily takes [a long] time"; i.e., "therapy does not necessarily take [a long] time." The latter statement will be recognized as rising out of client-centered therapy. Why is it that therapy does not take time? Because the client-centered therapist attempts by his behavior to create a "safe, accepting, non-threatening atmosphere," in which the only rules for client behavior consist of limits rather than specific directions. In terms of behavior theory:

If anxiety is a strong stimulus,
and
if a given instrumental (verbal) behavior is a cue for anxiety,
and
if the understanding and accepting behavior of the therapist

(*a*) rewards the instrumental behavior (higher-order reinforcement),
(*b*) reduces anxiety (second-order reinforcement),

then

(1) by gradient of reinforcement the immediately preceding covert and overt responses are reinforced,
(2) by generalization similar covert responses or response tendencies are reinforced which produces a decrement in their tendency to evoke anxiety.

We now assume that the instrumental behavior reinforced is less anxiety-producing than the immediately preceding covert response (defensiveness). This is consonant with the position of Dollard and Miller. Then the immediately preceding covert response is hypothesized to be (from gradient of reinforcement and drop in cue value for anxiety) prepotent over the verbal response just made. If not the first response would be repeated.

It follows from the above analysis that:

1. Progressive motivation is provided to express "dangerous thoughts," since those in the response sequence become progressively less "dangerous" than they were but retain their prepotence over other responses.

2. By generalization "similar" thoughts become "less dangerous," and since some of the responses similar to verbal response 1 are similar to verbal response 2, which in turn are similar to those of response 3, multiple sources of reinforcement and loss of cue value for anxiety of unexpressed thoughts "similar" to those of the expressed thoughts occur.

3. "The self-chosen pace of the client is a slow pace" is correct only in so far as the therapist arouses anxiety as well as reduces it, as Dollard and Miller point out but do not exploit.

At any rate, this analysis puts the issue just where it belongs. The two contradictory propositions, "therapy necessarily takes [a long] time" and "therapy does not necessarily take [a long] time," are both consistent with behavior theory. The contradictions in the propositions really reside in what is regarded as necessary behavior on the part of the therapist. This analysis in Dollard's and Miller's own terms shows that "dangerous" thoughts become less dangerous both directly and indirectly, and these thoughts should be expressed when the role of the therapist as an agent for anxiety is being minimized. Also, since the anxiety evoked by the unexpressed thoughts receives increments of extinction time after time, it may be seen that topics never mentioned or even "thought about" may be dealt with "adequately" in therapy, i.e., may become "conscious" and be dealt with in terms of anticipated consequences rather than reacted to indirectly by way of anxiety reduction (avoidance).

Finally, the problem of "time" in therapy has thus become an empirical problem, and the solution lies in the study of client or patient behavior in the presence of various constellations of therapist behavior. This conclusion is, of course, true only if behavior theory of the Dollard and Miller variety is considered to be a valid theory. For myself, I can see as much reason for accepting this brand of behavior theory as I can for accepting any current theories of psychotherapy and of personality. The test really lies in the study of therapeutic behavior, client and therapist, not in the learning theory. One unique function of learning theory can be to connect the apparently inconsistent or contradictory propositions of therapeutic theory and to direct attention to the consequences of therapist behavior, timewise and otherwise, in promoting adjustment. In other words learning theory will, I think, if properly used, emphasize the conditions of learning improved by the behavior of the therapist and suggest explanations of the resulting behavior of clients in terms of common processes.

As yet little has been done by learning theorists to relate behavior theory to psychotherapeutically based personality theories other than to equate the ego with the higher mental processes and the id with drives [2, 5]. Perhaps this has been because expressions such as "self" or "ego" seem to denote some kind of entity, some "little man up there." My clinical experience, however, leads me to believe that, when we clinicians speak of selves and egos or self-processes or ego-processes, we are speaking not of entities and not of processes considered as such but of the organization of the higher mental processes, of systems of processes, or unities which are open systems. The clinicians, in implying that these systems are open, mean that these systems are subject to change; that reorganization of the system may be going on more or less continuously and that in some sense disorganization may be present, especially when the change in the system is great.

Now, if the higher mental processes (thoughts and images) are considered responses, as Dollard and Miller so consider them, they must like all other responses be (or generate) stimuli. That is, they are response cues, and we may consider them as subject to generalization, extinction, etc.—that is to say, the propositions of behavior theory may be applied to them.

When we consider such functional stimuli and functional responses it is clear that they cannot be defined along physical or similarity dimensions like actual stimulus and response generalization. Furthermore the distinction between response and stimulus generalization which is possible in the non-functional definitions of the dimensions no longer seems warranted. Response generalization and stimulus generalization must mean the same thing, since for functional responses and functional stimuli the response is identically a stimulus. Furthermore, it would seem that since thoughts (functional cue-producing responses) are acquired by experience, by reinforcement, the generalization dimensions are themselves functional, i.e., are learned. For example, the inner response symbolized by "I am short" may connote "I am inadequate." The two responses are on the same generalization dimension or gradient. I think that this notion is entirely consistent with the position of Dollard and Miller, if, indeed, it is not part of their position.

If the notion that generalization of symbolic processes is based on experience, that the generalization dimensions are functional, is acceptable, then a given symbolic cue-producing response may lie on more than one generalization dimension. A thought may lead to several other thoughts which among themselves may be independent. Take, for example, a given thought, *A*, evoked by external stimuli, which is on three generalization dimensions, 1, 2, 3. Then *A* may evoke *B*, *C*, and *D*, which are on dimensions 1, 2, and 3, respectively. *B*, however, will not evoke *C* and *D* except by way of *A*. We may represent this situation by considering each generalization to be an arrow where the arrowhead represents the response highest on the gradient. Then *A* is at the point of intersection of the three arrows. It should be noted that the point of intersection need not be at the center of the arrows. *A* might be at a different point, might have a different value, on each of the three gradients. Assume now that response *E*, which is highest on gradient 1, is at the intersection of three more gradients and that one of these gradients is at the intersection of an additional three gradients, and so on *ad infinitum*. It is clear that we have a network in which the reinforcement or extinction of any response (thought) whatever has the possibility of reinforcing or extinguishing any response which is on a gradient in the network, provided that the sequence of responses leads to responses which are at the intersection of gradients.

It is not clear from Dollard's and Miller's discussion whether thoughts as responses are considered preparations for action. However, Sperry [*6*], on the basis of neuroanatomical considerations, so considers them, and such a view seems consistent with their theory. Considering thoughts as preparatory responses, then, it follows that a train of thought is a resultant of approach-avoidance gradients, and that which thought follows which depends on the resolution of conflict at the intersection of generalization gradients. The resolution of conflict at any point in the network would then depend upon the heights and slopes of approach-avoidance gradients to the number n, where n is the number of intersecting generalization dimensions. Predicting the next thought (response) at an intersection where $n = 10$ would obviously present certain complexities of analysis and probably could best be handled by considering the network as a whole in terms of a probability model.

If, in a general sense, the ego is considered the "executive of the personality," then the ego would, in the terms here used, be the largest network of response probabilities. I say largest because I am not neglecting the possibility of "split" personalities. In my opinion, this definition conforms to the clinical observation of unity in personality. The network is a *system* of response probabilities. It meets the notion that the ego is the conflict area of personality, and it meets the concept of "centrality" or of "consistency" in personality. The responses at the intersections of the gradients are certainly "central" in the sense that reinforcement or extinction, avoidance-of and approach-to the intersections, will have the most widespread and diffuse effects.

The response-cue character of thoughts and the notion of the collapse of stimulus and response generalization dimensions and of response sequences into single gradients or dimensions are also consistent with the concept of self. In terms of social learning and the course of socialization it seems that the most "central" thoughts, those at the intersection of the most gradients, would be self-concepts, thoughts which prepare the individual to respond to himself in given ways rather than others. The reinforcement and extinction of these thoughts would have the most widespread effects. Thus we might define such central thoughts or self-responses as constituting the "self-concept." These notions could be carried further, but I am not developing a theory of the ego and of the self; I am merely trying to indicate that, as is the case with psychotherapeutic propositions, the notions of therapy-based personality theories may also be translated to a single frame of reference.

The implications for psychotherapy of such a formulation seem to be obvious. If a person is neurotic or maladjusted, then these "central" thoughts (self-concepts) are, or tend to be, the cues for anxiety responses. Then all other thoughts connected with the central thoughts also become cues for anxiety, and thoughts which are quite different from the central thoughts and are on generalization dimensions removed from the central thoughts may come to evoke anxiety. And similarly for extinction.

Suppose now that anxiety is reduced by a typical instrumental behavior. Then all thoughts producing the anxiety through the generalization network become cues for the anxiety-reducing behavior, provided they are close enough temporally to the anxiety-reducing behavior. If the central thoughts are extinguished in ways which are described by Dollard and Miller, then all thoughts connected with them by generalization gradients may receive their increment of extinction, and hence the probability that the instrumental behavior will be evoked should be lowered.

Now we add the clinical observation that the most threatening behaviors tend to be those which call into question the "traits" and "roles" of an individual. Taking this observation together with the preceding development, we can perhaps glimpse how a theory of the ego or self would have implications for psychotherapy.

One implication is that "free association" is not at all free and that it tends to lead directly to central (self) thoughts; also that, if central thoughts are cues for anxiety, the procession of the response sequence arouses more and more anxiety when the central thoughts are anxiety producing with the consequence that avoidance gradients are then found to be above approach gradients. This is in agreement with nearly all theories of therapy.

Another implication is that a client may get closer to his central thoughts from the "periphery" of the network if he is in a therapeutic situation in which the external cues arousing anxiety are minimized; i.e., the tolerance of the client for self-evoked anxiety cues is maximized. It follows that more central thoughts (emotional problems) of the client can be "approached" before anxiety-reducing behavior (avoidance) is evoked.

This result is the same as the one reached earlier on a more limited basis; it is also more analytical.

Another implication is that, if the ego or self is a network, then the spread of reinforcement and extinction which occurs obviates the necessity of working through all central problems. Lowering of the cue value of thoughts for anxiety means that "internal courses of action" are not now as threatening as they were, and, therefore, that different courses of instrumental behavior (social interaction) will eventuate. In another paper [*1*] I advanced the notion that courses of instrumental, interpersonal behaviors, since they are largely anticipatory, tend to evoke consequences (response-wise) which reinforce that behavior. In the context of this paper, it is clear that the sequence of internal cue-producing responses is part of the sequence of instrumental behavior and therefore should be reinforced along with it, providing a double source of reinforcement for the internal responses. This explains why therapy can have such dramatic results; the anticipatory behavior of the client is influential in creating his own social (response) environment.

Finally, from the perspective that learning theory can give us, it seems that all this implies that we should look more and more closely at the role of the therapist in evoking and reducing anxiety reactions. If we do this with an open mind, I think that some of the contradictions, inevitabilities, and absolutes implied by the existence of so many psychotherapeutic systems will seem not to be so contradictory, inevitable, and absolute.

With respect to the contributions of the domain of learning to research in psychotherapy, it is my conviction that at present learning theory as such is able to contribute but little to actual research on the process of therapy. It seems to me that this is so because, as I stated earlier, translations or reductions of therapeutic theory to learning theory will not, in general, prove anything save that the sets of propositions are consistent or inconsistent, and the question of the actual denotata of the therapeutic propositions is not solved by the process of translation, although the translations may be richly suggestive. As far as I can see, the best use of the domain of learning can come from applying the experimental techniques invented in the laboratory to tap processes and abilities which are used in everyday life and thus may be substituted for the usual measures of social adjustment which are so unsatisfactory. Let me give an example of the type of research that may be used in studying psychotherapy, a type of research that to my knowledge has not yet been done.

It will be assumed first that maladjusted individuals may be divided into two classes, "repressive" and "vigilant." The repressive person tends to withdraw from threatening stimuli by "not noticing" them, by "forgetting" them, etc. The vigilant individual tends to "enter the stimulus field," to deal with the stimuli. He does not forget them. He keeps them in awareness, intellectualizes about them, and can and does talk about them. Lazarus *et al.* [4] have presented experimental evidence for the existence of such types and have shown that they can be identified clinically.

Assuming that our subjects, who are to undergo psychotherapy, have been identified, the following procedures are followed:

1. An association test, containing trait names or attributes drawn from psychotherapeutic protocols, is administered and scored for complex indicators. The hypothesis is that vigilant individuals will be differentiated from repressive by the exhibition of some extremely fast reaction times.

2. The words in the association test are administered tachistoscopically at speeds too high for correct recognition, with galvanic skin response and reaction time being used as criteria of subception. The hypothesis is that traumatic or anxiety-producing words can be identified through this procedure.

Clinical assessment, complex indicators, and subception are to be used to stratify the experimental population.

3. The words in the association test are then sorted by the subjects along subjective metrics of self and ideal. The hypothesis is that words with high scale discrepancies are consciously threatening. This has already been shown experimentally by Haigh [3].

Now, taking words which simultaneously (*a*) are complex indicators, (*b*) are associated with subception, (*c*) show high self-ideal scale discrepancies, we construct the following learning tasks: (1)

word mazes, inserting the "threatening" words at points favorable for recall and having the subjects learn the maze to a given criterion; (2) paired-associate lists, using nonsense syllables for the first lists and judicious admixtures of neutral and threatening words for the second lists.

Finally, repeat the above procedures after therapy.

The hypotheses are as follows:

1. Repressive individuals will forget the threatening words and remember the neutral words.
2. Vigilant individuals will forget neutral words close to the threatening words and will remember the threatening words.
3. After therapy both groups will be more efficient with respect to acquisition and retention. Of course suitable controls and corrections must be made for initial learning.
4. After therapy both groups will exhibit fewer threat reactions to a list comparable to the pretherapy list.

Such procedures, drawing upon already established experimental results, could be expected to show whether therapy had an effect upon tasks involving stimuli shown from the domains of learning, perception, and personality theory to be related to the conflicts of the individual. It should be noted also that such tasks have a certain resemblance to everyday learning tasks which require serial learning, associative learning, and accuracy of perception under conditions of ambiguity.

To sum up, it has been my purpose in this discussion to indicate, from a clinician's standpoint, contributions to the theory and practice of psychotherapy that might accrue from the translation of psychotherapeutic propositions to those of learning theory. The main advantage theorywise appears to be the possible reduction of disjoint propositions to a common frame of reference to the effect that new possibilities may appear and be explored systematically. The main contribution of the techniques of investigation involved in the learning laboratory seems to be confined at present to the objective study of the outcomes of psychotherapy.

REFERENCES

1. Butler, John M., The interaction of client and therapist, *J. abnorm. soc. Psychol.*, 1952, *47*, 366–378.
2. Dollard, John, and Miller, Neal E., *Personality and psychotherapy*, New York, McGraw-Hill, 1950.

3. Haigh, Gerard V., The role of value and threat in perceptual orientation, unpublished Ph.D. dissertation, University of Chicago, 1951.
4. Lazarus, Richard S., Shaffer, G. Wilson, Fonda, Charles P., and Heistad, Gordon T., Clinical dynamics and auditory perception, *Amer. Psychologist*, 1950, *5*, 305–306 (abstract).
5. Mowrer, O. H., The law of effect and ego psychology, *Psychol. Rev.*, 1946, *53*, 321–334.
6. Sperry, Roger W., The mind-brain problem, *Amer. Sci.*, 1952, *40*, 291–312.

Learning: an Aspect of Personality Development

DONALD SNYGG

I must confess that when I came here I was worried about how some of the things I have to say would be received. When I began to consider the topic of this symposium some of the things I thought of seemed so extreme and so bad tempered even to me that I wondered if they could really be true. I am accordingly very grateful to the other participants for the reassurance I have received from their remarks. I am particularly grateful to Dr. Spence for his disarmingly frank and modest statement that his learning theory at the present time has no application to any but the most simple behavior. The other participants seem to be in substantial agreement, not only about Dr. Spence's theory but also about learning theories in general.

I suppose I should be happy about this situation because it removes much of what I have to say from the realm of controversy and consequently from the risk of emotion and hard feelings, but as a psychologist I find it somewhat depressing.

From any practical point of view the basic problem of psychology is the problem of learning. Most psychologists are paid to help people to learn. Whether we are academic psychologists, educational psychologists, clinical psychologists, or industrial psychologists we are supposed to be experts in learning. This is embarrassing because the truth is that nobody knows very much about learning. Hilgard, writing in 1948, sixty-two years after Ebbinghaus, said, "There are no laws of learning that can be taught with confidence. Even the more obvious facts . . . are matters of theoretical dispute" [3, p. 326]. In psychology theories of learning are taught and argued about but not used. The truth is that the main advantage the practicing psychologist has over an intelligent layman in dealing with learning is his greater experience.

The inadequacy of our current learning theories for practical purposes is clearly revealed by their failure to have any effect on educational practices and objectives. Only a few years ago the psychology of learning constituted the core of educational psychology. Now learning has been pushed far into the background by mental hygiene and personality. Although a few teachers try to give themselves an air of authority by talking about conditioning when they mean learning, the practices of the typical school have remained completely unaffected by any learning theory of the last 35 years.

This seems strange when we remember that learning first became an object of psychological research because of the practical needs of teachers. As a psychological concept, learning is rather new. Baldwin's *Dictionary of psychology*, published in 1902, did not even list the term, and Warren's 1934 dictionary referred it to applied psychology only [*1*]. Learning, as English has pointed out [*1*], is an educational concept. It was no accident that compulsory education laws were soon followed by Thorndike. Confronted for the first time with large numbers of pupils who were not prepared to learn what the teachers were prepared to teach, educators turned to psychology for help.

Many of them are still looking. The rest have given up.

The sad truth is that, after 50 years of careful and honest and occasionally brilliant research on the nature of learning, the only people who can be proved to have received any practical benefits from learning theory are the learning theorists themselves. The very inconclusiveness and complicated nature of our current learning theories, which make them useless to applied workers, have proved to be occupational assets to the learning specialists. They can, if they wish, make rather good professional careers out of attacking the weak points in one another's theories, much like the shipwrecked Scotsmen who made a good living by taking in one another's washing.

My gibe at the complicated nature of current learning theory may seem unjust. It is not the fault of the theorists, some may feel, if their theories become too complicated for teachers to understand or for applied psychologists to use. I am not so sure. Teachers are not particularly stupid, and most really productive ideas in the history of science have been fairly simple. Many theories are complicated because a great many qualifications have had to be added to patch up an idea or a conceptual scheme which was not very good in the first place. There are clear signs that our psychologies of learning have reached this futile stage of patching what had better be thrown away. Harlow [2, p. 27] has had something to say about this. Writing about learning theory, he comments, "A strong case can be made for the proposition that the importance of the psychological problems studied during the last 15 years has decreased as a negatively accelerated function approaching an asymptote of complete indifference."

Of course, I may be wrong. It is just possible that one of our presently inadequate theories may, in the future, be developed into something really effective. But I don't think so. The kind of generalization that most learning theorists are using is not the kind of generalization which can be applied to the behavior of actual people in social situations. Somewhere our analysis of learning has taken a wrong turn. I am inclined to agree with English [*1*] that the mistake was made when it was first assumed that learning is a relatively independent psychological process.

There are very good reasons for questioning the usefulness of the "process" concept of psychology itself, but Krech has protested so effectively [*4*] against this particular way of fragmenting psychology and people that I do not need to burden you with my version. The important thing is that, as English pointed out in his 1951 address to the Division of Educational Psychology [*1*], the phenomena which we are accustomed to group in the category of "learning" phenomena are lumped together for practical reasons, not for systematic or theoretical reasons. Learning is always defined in terms of practical results. To prove learning we have to prove increased efficiency in behavior.

"Before psychology took over, the verb 'to learn' was used in common speech to designate the whole complex process of reorganizing, improving, adjusting one's behavior" [*1*, p. 328]. When psychologists took up the study of learning they had no safe grounds for making either one of two assumptions that underlie all theories about the "learning process."

The first assumption is that all learning, i.e., all improvements in behavior which cannot be ascribed to maturation, is the result of a single psychological process. The second assumption is that this hypothetical process is not the cause, except indirectly, of any other changes in behavior. If either one of these assumptions is not true the postulate that learning is a single psychological process cannot be valid.

Actually, anyone who tries to deal with learning in relation to other aspects of human behavior must eventually give up one or the other of these assumptions. Psychologists who begin by equating learning with the process of association are bound, sooner or later, to run into "learning" behavior which does not fit the association pattern and to conclude that learning is not one process but two or more. Those of us, on the other hand, who take an open dynamic field as the model arena for learning behavior are sure to find the model adequate for dealing with other aspects of behavior and to conclude that learning is not an independent process at all but simply another manifestation of field organization. After contending for some years that learning is not two processes but one I have just come to realize, with the help of a few nudges from English [*1*], that what I have really meant is that learning is not a separate or unique psychological process at all.

Be that as it may, the assumption that learning is a more or less independent psychological process has had a definite and, I believe, unfortunate effect on learning research and theory.

In the first place, if we believe that learning is a process, it is quite natural for us to begin to think of learning as somehow independent of the people who learn. The time is past when psychologists were looking for the "true" form of the learning curve. But it is still standard practice to obscure the individual nature of behavior and learning by combining the data from several subjects.

In fact, if learning is a process it is not necessary to study people at all. White rats are much more prolific than people, are cheaper to feed and house, and are usually willing to work for room and board. Do not misunderstand me. I like white rats. When properly handled they are gentle and often affectionate. Some white rats are better company than some people. I have a high respect for their intelligence. But I should like to say that, whereas all psychologists are properly aware of the danger of assuming that rats are just like people, many of them do not seem sufficiently aware of the dangers involved in assuming that people are just like rats.

If we assume that learning is a process and that it is the task of the learning psychologist to study that process it does seem to follow that the place to study it is under the simplest possible conditions, probably best represented by a white rat in a single section T-maze. No matter what apparatus is used it is sure to be one which limits the possible behavior of the subject to two simple alternatives at each choice point. This is done to simplify and facilitate record keeping. Then, so that the animal will not learn too fast, all but one of the possible cues that it might use for solving the problem are eliminated. Personality differences are further minimized by arranging that all subjects share a common physiological tension, usually hunger, during the experiment.

Such practices would make no sense if the experimenters were trying to learn about people or even about rats. But they follow quite logically if the experimenter is trying to study a process. Since the behavior of the subject in most experiments is pretty well limited by the nature of the apparatus that is used, what we may actually be doing is studying the limitations of our apparatus.

At any rate we have reached something that looks like a dead end. It may not be. I have been wrong before and expect to be again. But just in case it is a dead end we should be calling for volunteers to go back and try to find another approach.

Anyone who tries to do this will be in new territory and will have to go where there are no maps. In a science as new as psychology we cannot afford to overlook any possibilities, and I hope that a number of possible approaches will be explored. But if I had to take my choice I would go back and start where the early "process" theorists did—with the practical problems of the people who are hired to help others to behave more effectively.

If we do so we shall find that this group and its problems have changed a great deal since learning psychology left it and went to the cats, rats, and dogs 50 years ago. For one thing the group now includes therapists as well as teachers; for another, the teachers are dealing with problems that are different from those that engaged them then. Teachers are now much more aware of their responsibility for the development of citizenship, character, and effective personality than they were 50 years ago, when it was generally assumed that whoever was well informed would also be good.

It is too bad that psychologists became interested in learning before this happened. By setting up our definition of the learning "process" too soon we have cut the learning psychologist off from the major problem of our schools and left it in the exclusive possession of the mental hygienist. And by making a dichotomy between the learning of facts and the development of personal character and attitudes we have encouraged the teacher to make a similar distinction and to concentrate on one or the other. It was several years ago that a teacher told me that character education was awfully important and that she intended to do something about it sometime. She would, too, if she could just get all her pupils through the multiplication table in time to do it. But I think it could happen today.

Now how can we repair this fracture and get the changes in an individual's general style of behavior which we have called personality development related to the changes in his specific acts which we have called learning? Stated in that way, there is an obvious relationship. If you are a behaviorist the acts are apt to be considered the result of the formation of separate S–R connections, and personality change is visualized as being nothing more than the sum total of these individual learnings, each one of which is essentially unrelated to the others.

I don't like this because it does not seem to give sufficient weight to the fact that in a situation where an individual is free to choose among a number of alternatives his behavior will show a marked degree of individuality and consistency. His learnings will conform to his personality pattern at the same time that they are changing it.

The following report, written by a college student, is an example of the sort of thing I am talking about:

> During my years in grade school I was always at the top of the class. I always got very high grades in all subjects but especially in arithmetic and silent reading. I spent a good deal of school time reading encyclopedias for fun and scored 100 in the eighth-grade arithmetic Regents' examination while I was still in the seventh grade. I also won the county spelling contest that year. By passing the Regents' exams during my seventh year I was ready to go to high school, but during the summer I suddenly decided that I didn't want to go to high school and my parents couldn't understand the reason.
>
> It was simple. My parents were *strict* members of a very strict religious sect. I would rather face a firing squad at dawn than face my father if I had "sinned." That summer my mother somehow found out that I was masturbating. She pointed out the village idiot and explained that she got that way by masturbating. She also threatened to tell my father.
>
> I decided that I had ruined my brain and could never learn anything more. I was afraid that I would flunk high school courses and my father would find out the horrible truth. I started to school with that idea in mind and barely scraped through the first year with the exception of math. I flunked that twice.
>
> At the beginning of the second year an intelligence test was administered. The students were not allowed to know the results, but a boy who was working in the NYA program told me that I had a score of 90 which was about average. My work began to improve a little. Later that year one of the boys got a book on sex, and "eminent authorities" stated that masturbation did not lead to insanity or mental deterioration. I graduated from high school fourth in the class.
>
> When I entered the army I scored 145 on the Army General Classification test, and in three verbal intelligence tests since then I have consistently scored high (not under 135).
>
> For two years the idea that I was stupid and had ruined my brain kept my school grades low and even depressed my IQ thirty to forty points below my usual score.

If we think of learning theory and personality theory as independent areas of psychology we are at a loss when we have to deal with situations like this. Learning theories are set up to explain why people change, and personality theories are set up to explain why they don't change. As a result personality theory is incompatible with learning theory, and we cannot work out any relationship between them.

When we are watching an instrumentally depersonalized animal in a T-maze or problem box this is not likely to seem important. But when the task is to predict the behavior of an individual who is free to pursue any of a number of alternative courses of action it becomes obvious that neither type of theory is adequate by itself. And since they are based on conflicting premises and lead to conflicting conclusions we can't use them together.

The conceptual scheme which I think would avoid this compartmentalization has been described elsewhere [5, 6], and I do not have the time to describe it in detail here. But if you can remember the boy who "ruined his brain" I should like to call your attention to three things.

The first is the catastrophic effect that one item of information or, rather, misinformation, had on the child's personality pattern. This personality change is quite obviously not the result of the mere accretion of one more fact.

The second is that the kind of learning of which he was thereafter capable was limited and determined by his view of the situation. As a good student he had excelled in arithmetic; as a boy who had ruined his brain he failed algebra twice.

The third is the way later items of information acted as precipitating agents to throw the personality into a third pattern in which he was again an able student.

Here we have the main characteristics of a dynamic field operating on something like all-or-none principles with change by transformation rather than by addition. The parts of the field are so interdependent that they are very resistant to change. But when change is possible and does occur all parts of the field are affected. If we adopt a dynamic field as the model for our conceptual system it is easy to avoid the separation between learning theory and personality theory that has caused so much trouble. Let us, for instance, assume that all behavior is determined by the behaver's perceptual field at the instant of action. This is a field theory, since the perceptual field is an organized field with space-time dimensions which, like all dynamic fields, tends to remain organized in the face of external interference. From this point of view, the consistency, relatedness, and stability of an individual's behavior, which constitute the problem of personality psychology, are the natural results of the organized nature of the causal field. An individual shows a particular style of behavior as long as his perceptual field maintains a particular pattern of organization. At the same time all the specific acts of the individual, changes in which constitute the problem area of learning psychology, are simply the ways in which the organization and integrity of the perceptual field are maintained. We avoid pain, move from the hot sun to the shade, seek food when we are hungry, and rest when we are tired to maintain the balance between the self and not-self in the perceptual field. These acts could also be considered as resulting from aspects of homeostasis in a physiological field or as physiological drives.

But we do other things which cannot be considered to have homeostatic or physiological value without seriously changing the meaning of those terms. I am referring to such things as martyrdom and suicide and writing books when no profit is expected and the actions of Harlow's monkey in putting in a 19-hour day keeping a window shade up. The perceptual field, as an organized entity, includes a perception of the future, at least of its existence. Maintenance of a satisfactory relationship between the perceptual self and this vaguely perceived future requires unceasing vigilance and activity. The future is uncertain. Since it is uncertain we can never gain a permanently satisfactory relationship with that part of the field. Under the threat of the future we can maintain the organization of the perceptual field only by striving for better and better relations between our perceptual selves and the rest of the universe through an increased feeling of worth, value, power, and acceptability.

From this point of view the primary goal of all behavior is the achievement of a more adequate self. Personality development is the goal of all our acts, and learning is the result of our attempts to achieve it.

Can this theory, which does not give any particular priority to the so-called physiological drives, be applied to the behavior of rats? I think so. Although a rat's concept of the future is probably even vaguer than our own, any animal with distance perception can be assumed to have some perception of time [6, p. 350]. Many of the latent learning experiments have shown learning without apparent tissue-tension motives, and I have seen starving rats give up food and water in order to keep possession of an activity wheel from their cage mates. The theory certainly fits the behavior of Harlow's monkeys [2].

It seems to me that this conceptual scheme has a number of advantages. On a theoretical basis it provides a much better framework for predicting behavior in a free situation than any of the specialized learning theories because the personality of the learner, in the form of his perceptual field, is an integral part of the conceptual scheme. If our investigations of learning behavior are to be helpful to the teachers, psychologists, and other people who are interested in assisting personal development, we will have to stop thinking of learning as an independent process and shift to some kind of unified-field theory. The one I have sketched happens to be the one I like.

Whether you like that particular scheme or not, I hope that you will consider the disadvantages of the present isolation of learning theory which I have tried to point out. Whatever conceptual scheme we use, psychology, education, and psychiatry could all benefit from a greater use of human subjects in research, with animal subjects being reserved primarily to check the validity of hypotheses developed from the observation of human subjects in situations where there are enough courses of action to give individual patterns and idiosyncrasies a chance to show.

If we do this I am confident that we shall find that the learner is not the passive victim of his environment but an active explorer and creator of his own world. He is not a puppet at the mercy of the stimuli which bombard him or even of his own hunger pangs. In perceiving, in learning, in forgetting, in imagining, in rationalizing, he is selecting from all the potential aspects of his world those which best satisfy his need for personal growth and development. If we forget the active and purposive role which the learner plays in the achievement of his own personality and his own future we are bound to deal with him arbitrarily and unrealistically and to frustrate him and ourselves.

REFERENCES

1. English, H. B., Learning—"they ain't no such animal," *J. educ. Psychol.*, 1952, *43*, 321–330.
2. Harlow, H. F., Mice, monkeys, men, and motives, *Psychol. Rev.*, 1953, *60*, 23–32.
3. Hilgard, E. R., *Theories of learning*, New York, Appleton-Century-Crofts, 1948.
4. Krech, D., Notes toward a psychological theory, *J. Pers.*, 1949, *18*, 66–87.
5. Snygg, D., The psychological basis of human values, in Ward, A. D. (Ed.), *Goals of economic life*, New York, Harper and Brothers, 1953.
6. Snygg, D., and Combs, A. W., *Individual behavior*, New York, Harper and Brothers, 1949.

"Errors": Theory and Measurement

R. B. AMMONS

I should like to start off by indicating briefly what I hope to cover in this paper. The major sections will deal with the following topics:

First, *"theoretical" versus "antitheoretical" approaches to the understanding of behavior.* This section concerns itself with a suggestion as to a distinction in the use of the word "theory," and with an indirect analysis of why people are often opposed to rigorous theorizing about the behavior of organisms.

Second, *part of a theory or postulate system dealing with "errors."* The phenomena with which the theory deals are some of those mentioned by Freud [2] in *The psychopathology of everyday life,* while most of the basic concepts are derived from learning theory. The "complete" theory will not be presented here because of time restrictions.

Third, *a brief report of an exploratory experiment designed to test roughly one of the deductions from the theory.* In this experiment students, taking a final examination of the essay type, were given relatively hard or relatively easy questions. "Errors" were measured in terms of number and type of writing deviations.

Fourth and finally, *some suggestions as to situations lending themselves to manipulations of the type necessary to test and extend the theory.*

What about the value of theories in psychology? Are they performing an intellectual service, or are they premature and otherwise undesirable? In the last few years there has been a series of attacks on theories and theory construction in psychology. It is my impression that by far the greatest number of comments on *rigorous* theorizing are negative. For this reason, I was particularly pleased at receiving an invitation to participate in this symposium, and to have an opportunity to discuss this apparent trend. I should like to raise a small voice against the antitheoretical points of view.

First, I should like to suggest a useful convention in talking about theory. When I use the term "theory," I am referring to a rigorous postulate system, the simplest form of which is the Aristotelian syllogism. Any formulation short of this rigor, I prefer to call speculation, without, however, in any way reflecting on the potential validity or utility of the ideas. Thus Guthrie's and Tolman's learning theories would not be called theories, but complex speculation about possible theorizing in psychology. E. J. Gibson's postulate system [3] dealing with the effects of generalization and differentiation on verbal learning would be called a theory. Hull's italicized statements in *Principles of behavior* come close to being theory in this sense.

This convention is useful and, I believe, should be adhered to for its psychological effect. I imagine that each of us here has had the experience of speculating about a phenomenon, and feeling that his speculations account for it quite adequately, only to find that when these speculations were reduced in writing to a minimum set of statements there was some "fatal" inherent contradiction or glaring hole in the reasoning. By carefully distinguishing between the loose and the logically precise system, between speculation and theory, one reminds oneself of the long step between them. The *ad hoc* loose system such as that of the early Gestalters [5] is capable of predicting everything or nothing, whereas the phenomena predicted by the *Mathematico-deductive theory of rote learning* [4] are usually critically testable.

I recently had the interesting experience of attempting to teach a highly selected group of graduate students how to go about formulating simple theories. In fact, we spent a whole semester working on this problem. As a clinician, I was fascinated by the responses of these students to the idea of constructing theories and tried to make observations as to the possible bases of these responses. Let us make the assumption that the present widespread attack on theories and theorizing in psychology is merely the organized expression of attitudes and feelings widespread in our culture. Then by observing my students I would be able to come to some sort of tentative conclusions about the reasons for the attack.

What difficulties *did* the students have, and what were their objections to rigorous theorizing?

The first problem was comprehension. A postulate system is hard for most students to understand. They say, "Why not put it in *simple* language?" Now, an examination of this comment shows that it is somehow at variance with the facts. The simplest, most concise way to express a complex set of relationships in words is to set up a postulate system. Any attempt to substitute literary language and exposition for the postulate system leads to the use of large numbers of words and a great decrease in precision, especially in regard to implications. Why, then, is the system hard to understand? My students discussed this point at considerable length and concluded that the difficulty for them was the result of inadequate training, particularly in mathematics. They felt that they simply had not been taught to think clearly and simply. This is all very well, of course, but *how* are we to persuade students to take mathematics? How many of us who serve as advisors are willing to "go back" and make up our own deficiencies in logic and mathematics? The average amount of training of psychologists in symbolic logic and mathematics is, I am afraid, appallingly small.

The second objection to postulate systems was that they call for operational definitions of terms, and when we define terms this way *we leave out meanings.* That is, the words lose some of their glorious essence. I call this a "something-more-than" attitude and have pointed out to persistent objectors along this line that they might too be called "something-more-thaners." This attitude grows from a satisfaction with things as they are and a failure to realize the nature of language. I have often pointed out that the meanings of terms are not inherent, but are given by experience. By setting up methods for more sharply defining words we are bound to step on somebody's toes. Interestingly enough, it is most often the private, emotionalized meaning which is left out in this process, and which the "something-more-thaner" tries to force us to include.

A third, and related, objection to logical systems is that they are incomplete and artificial. They leave out much that is obviously present in the world, and lead to the ignoring of important facets of the phenomena. Here again is the "something-more-than" feeling rearing its ugly head. I like to point out to the student at this point that there is nothing to keep him from expanding the system to include those things which he feels have been neglected. He often counters that they could never *all* be put in, so the system cannot have any great value and may even mislead by oversimplification. I have found that careful investigation of the student's feelings usually reveals a real resistance to thinking carefully and precisely about the phenomena because of some kind of often hidden, personal needs. That is, he has some peculiar personal reservations about dealing with this sort of thing coldly and intellectually.

The fourth objection is in some ways similar to the third. It takes the form, "We don't know enough to set up a postulate system about *that.*" This objection often seems to grow out of a desire to avoid the rigors of careful logical thinking. The student would prefer to go on carrying out relatively unrelated experiments to obtain empirical results *ad infinitum.* This is the Baconian approach, of course, and perhaps grows from a need to hoard relatively unorganized information. To this objection, I usually point out that we never will know *everything* about *anything* and that the outcome of adopting this approach might well be a permanent avoidance of theorizing. Several students have then pointed out that "premature" theorizing, or theorizing before we have "enough" data, may lead to a blindfolding of future research. I am quite willing to agree that this has happened and probably will continue to happen. However, the person who allows a theory to restrict his thinking, or who becomes violently partisan in an attempt to "prove" the theory, would probably be just as dangerous if there were no theory. He would then probably feel compelled to defend some particular finding or findings in the same sort of blind partisan way. But let's not blame the theory for the scientist's personal shortcomings.

In general, there are some obvious reasons why the student, and for that matter almost anyone, might and often does feel antitheoretical. Constructing or even comprehending a "simple" postulate system is hard work and is often not immediately rewarding. Then, when we have a system worked out, it ordinarily calls for continuous revision and extension—more work. For the system to be of real value, there must be a great deal of systematic work done, and we Americans seem to find it hard to work systematically for long periods of time. Our individualism also often keeps us from working on problems which have the misfortune to have been suggested by someone else. Finally, there is a real anti-intellectualist feeling on the part of many "clinicians" who, probably rightly, find an intuitive approach most valuable in *practicing* their art, and then generalize this into opposition to "sterile, meaningless theorizing."

Well, I might say that trying to teach my class something about theorizing was a stimulating experience, but I uncovered no particular talent for theorizing.

Let us turn now to the principal topics of this paper, a tentative theory of errors, a report of some related research, and some proposals as to how to study the situation experimentally. I should like to start out by defining in a tentative way the terms used in the theory.

Error: A response other than that appropriate to the motor set present, where this response is appropriate to other parts of the stimulus complex.

Response: Observable striated muscular behavior by the individual.

Motor set: Bodily orientation for the performance of a given behavior, inferred jointly from the instructions given by the experimenter or subject to himself and the physical orientation of the person. We can to some extent get at it by asking the subject what he intends or intended to do, or by setting up an objective criterion for determining whether or not the physical orientation would allow the performance of the task.

Appropriate response: The response which the individual says he intends or intended to make and for which he is physically oriented is the appropriate response to the motor set. Appropriate responses to other parts of the stimulus complex are those which would be most frequently made if those parts of the stimulus complex were dominant.

Stimulus complex: Various components which make up the stimulus such as stimuli from motor set, specific drive stimuli, and external stimuli. Any of these can be changed relatively independently, changing the stimulus complex.

Dominance of a component of the stimulus complex: A drive stimulus is more dominant as the drive becomes stronger. When the subject is asked to describe a situation, a particular stimulus component is dominant to the extent that it is mentioned earlier in his description. Frequently this dominance must be inferred from the past history of the individual. The report may not be accurate from the point of view of the experimenter, as in the case of the individual who has always hated a sibling and now reports that his emotion is one of love and affection, yet behaves as if he still hated her. This is admittedly shaky ground for inference in some cases.

Drive stimuli: Those stimuli characteristically noted by the human organism in connection with hunger, thirst, sex frustration, fear, anxiety, etc. One could infer the presence of such stimuli in terms of strength of drive.

External stimuli: Environmental energies which affect the receptors of the organism. When the organism is oriented in such a way that the receptors can be affected by the energy and the energy is sufficient to stimulate the receptors, stimulation is normally assumed to take place.

Strength of the response tendency: Latency of the response, physical strength of the response, and probability of the response occurring in the presence of or closely following the presence of a given stimulus complex.

Stimulus similarity: Stimulus complexes are similar to the degree that they contain similar components and are relatively less separated along the various discriminable continua.

Strength of drive: Might be the self-rating of the individual or might be inferred from the past history of the individual with respect to the time since drinking, time since eating, number of times a pleasant or unpleasant consequence has followed a particular stimulus complex, etc. Thus drive stimuli can be associated with primary or secondary drives as conceived of by Hull. Emotions are considered to be drives.

Reward: The satisfaction of some need, goal-object consumption, or avoidance of noxious stimulation.

Thus we have the basic concepts: error, response, motor set, *appropriate* response, stimulus complex, *dominance of a component* of the stimulus complex, drive stimuli, external stimuli, strength of response tendency, stimulus similarity, strength of drive, and reward. Although not all of these are used in the part of the theory which I am going to outline today, I thought it might provide a better overview to include them.

Only the first four of eleven postulates will be presented, along with several deductions generated by them.* It will be seen that I am not doing anything particularly original, just attempting to formulate some of Freud's ideas in terms of concepts from learning theory. Many of the approaches taken here have been suggested by other writers such as Mowrer [7] and Miller [6]. I consider the present form of the system to be tentative.

The whole system is based upon the assumption that behavior is predictable. The first postulate reads as follows:

> POSTULATE 1. *To any stimulus component or complex, there are a number of possible responses. The strengths of the response tendencies differ. Thus there is present a "strength" hierarchy of responses to any given stimulus component or complex.*

It should be noted that a response may produce stimuli for other responses.

> POSTULATE 2. *The more similar a stimulus component or complex is to another given stimulus component or complex which*

* An outline of the "complete" system can be obtained by writing to the author.

has regularly elicited a response in the past, the stronger the response of this kind now elicited by the new stimulus.

This implies a stimulus generalization gradient. In a free association situation, the stimulus includes the feeling present, intellectual content, traces of these from previous recent verbalizations, and present physical surroundings. Concepts are all related; stimuli are all related.

POSTULATE 3. *The stronger the drive, the stronger the response.*

This postulate will interact in effect with Postulate 2, since drive has a characteristic stimulus value which will be changed by increases or decreases in the drive.

Deduction 3a. The more drive present, the less similar the external stimulus need be to the stimulus which in the past has regularly elicited a response, for it to be elicited with the same strength.

In a free-association situation, a very pressing drive will "force" a response to relatively non-pertinent stimuli.

POSTULATE 4. *The components of a given stimulus complex may in isolation elicit different responses. When the components are combined in the stimulus complex, the greater the dominance of a given component and the greater the strength of a given response tendency associated with it, the more likely the stimulus complex is to elicit this response.*

Deduction 4a. If a response has been regularly elicited under a low drive and is now elicited with a high drive of the same kind present, we will observe an increase in "errors," providing the strongest response tendencies to the motor set and the drive are different and that to the motor set is dominant.

Deduction 4b. If a response has been regularly elicited under one drive, and the drive is changed to another without altering the other stimulus components (especially motor set), there will be more errors, providing the appropriate dominant response to the drive-stimulus component from the original drive was the same as that to the motor set, but that to the new drive stimulus is different from that to the motor set, the motor set staying the same.

> *Deduction 4c.* To the extent that a single stimulus component dominates the total stimulus complex, the successive responses given by an individual will be more similar to each other.

Strong emotion leads to stereotypy of responses, as does instruction-induced "motor set," and the "same" physical stimulation. In free association, problem areas will be talked about more frequently than other areas. In the case of errors, we find that certain kinds are quite frequent, i.e., certain types of slips of the tongue and certain kinds of accidents in the accident-prone person. These errors should indicate the life areas in which the person has problems and thus be of diagnostic value to the clinician.

> *Deduction 4d.* Other stimulus conditions being approximately equal, if one arouses feeling about an error, he should get real-life responses associated with a similar set, emotion, or drive more quickly than if no feeling is aroused.

If one arouses guilt feelings, for example, by calling attention to an error, these will tend to become relatively dominant in the stimulus complex and will form a stimulus complex increasingly similar to that real-life situation in which the error was originally made as a response.

So much, then, for the first few postulates and deductions. Having made some predictions, naturally we wanted to test one or more. A likely situation presented itself. The writer was teaching a class in history and systems of psychology, using essay-type examinations. Two sets of four questions were made up for the final examination. In both sets, questions 1 and 3 were the same, whereas 2 and 4 were *much* harder in one set than in the other, covering topics not mentioned in class and only indirectly in the textbooks. Thus we had two relatively easy control questions for each group and two other questions, either relatively easy or very hard. The difficulty of the questions was checked by asking the students to rate the questions after the examination. Their ratings agreed with our assumptions.

The questions were mimeographed on separate pieces of paper, one question to a page, and booklets made up of them, in the order 1, 2, 3, and 4. The two resulting sets of booklets ("easy" and "hard") were arranged in one random order, and passed out in this order on the day of the examination. Students were told that they could write on both sides of a page, but should write only in the booklets, and only during the time allotted for a particular question, which would be 12 minutes. There was no observable difficulty in carrying out these instructions. A total of 36 students completed the test and the experiment, 18 in the "easy" and 18 in the "hard" group.

This experiment provides a rough test of Deduction 4*a*. The dominant motor set calls for the response of writing answers clearly and accurately—i.e., for "good" examination behavior. Students feel, during such examinations, numerous emotions which could lead to contradictory behavior, feelings such as fear, anxiety, and aggression. The introduction of difficult questions, our questions 2 and 4 in the "hard" series, tends to intensify these feelings, as indicated by the students' own ratings of the questions. Thus we have the dominant motor set and the increased drive which would lead to different behavior if appropriately acted out, so that the initial conditions of the deduction are present. The prediction is that "errors" will increase differentially in the case of the "hard" questions.

To test the prediction, we counted as "errors" all erasures, cross-outs, write-overs, and misspellings in each answer by each subject. The variance of these scores was then analyzed,* following a procedure suggested by Block, Levine, and McNemar [*1*]. In effect the interaction between questions and groups is tested over a residual error term based on pooled within-groups interaction between questions and individuals. Our "easy" group scored totals of 108, 102, 116, and 88, while our "hard" group scored 62, 83, 70, and 78 on the four questions. It will be noted that the group patterns differed in that the "hard" group made relatively more errors on the hard questions. This interaction was significant at beyond the 1 per cent level, while the difference between groups did not reach the 5 per cent level. Although the prediction and the outcome agree, one should be cautious in interpretation, since many errors can creep into experimentation of this rather complex kind.

We have done a good deal of thinking about other experimental situations which might be used in the testing of deductions from the system. One might ask typists in training to copy emotional and neutral selections for practice. Subjects could be asked to write in longhand or shorthand about neutral and loaded topics. They might be asked to copy a story while studying a TAT card or other similar pictorial material. Many possibilities remain to be tried out.

To recapitulate briefly, in this paper I have tried to do several things: (*a*) to show that theorizing is a defensible procedure; (*b*) to

* I particularly wish to express my appreciation to Mrs. Sylvia Post, who became much interested in the study and generously gave of her time in performing most of the statistical calculations.

present a preliminary formulation of a theory of errors; and (*c*) to show how the variables might be measured experimentally.

REFERENCES

1. Block, J., Levine, L., and McNemar, Q., Testing for the existence of psychometric patterns, *J. abnorm. soc. Psychol.*, 1951, *46*, 356–359.
2. Freud, S., *Psychopathology of everyday life,* in Brill, A. A. (Tr.), *The basic writings of Sigmund Freud,* New York, The Modern Library, Random House, 1938.
3. Gibson, E. J., A systematic application of the concepts of generalization and differentiation to verbal learning, *Psychol. Rev.*, 1940, *47,* 196–229.
4. Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., Perkins, D. T., and Fitch, F. B., *Mathematico-deductive theory of rote learning,* New Haven, Yale University Press, 1940.
5. Koffka, K., *Principles of Gestalt psychology,* New York, Harcourt, Brace, 1935.
6. Miller, N. E., Theory and experiment relating psychoanalytic displacement to stimulus-response generalization, *J. abnorm. soc. Psychol.*, 1948, *43*, 155–178.
7. Mowrer, O. H., *Learning theory and personality dynamics*, New York, Ronald Press, 1950.

Some Current Research Issues in Clinical Psychology

J. R. WITTENBORN

For the most part it may be said that clinical psychologists serve individuals in two different respects: one, *diagnostic;* the other, *therapeutic.*

Diagnosis comprises several functions: describing the patient, relating the patient's present condition with circumstances of the past, trying to understand the patient in terms of possible relationships between the past and the present, and, finally, applying this understanding in some manner to anticipate the course of the patient's development. The tools for diagnosis, tests, current observations, and information from the case history of the patient have been emphasized in our research and training; the inferential manner in which these tools are used in order to gain an understanding of the patient has received little attention. Although anticipation of the subsequent development of the patient requires deductive application of inferences, the rules and principles for drawing our inferences and for conducting our deductions are not explicit and are guided by no consistent theory.

The purpose of the *therapeutic* process is easy to describe: therapy is designed to *change* the patient in one or more respects. Since many different kinds of behavioral changes are sought for patients, some difficulty in describing the method of therapy can be expected. This difficulty is compounded by the fact that there are, in effect, countless methods whereby clinicians seek such changes.

Since the diagnostic function we have described can be construed as an attempt to understand how the patient changed from his original state to his present one and to anticipate how he may further change, and since the therapeutic function is defined as an effort to produce certain kinds of changes, we may say that the clinical psychologist is concerned with changes in human behavior. Unfortunately, as soon as we see the clinical psychologist as concerned with changes in human behavior, we become aware of two weaknesses:. (1) aside from a few esoteric laboratory aspects, our understanding of how to produce changes in human behavior is so obscure and incomplete as to be contradictory in its implications and possibly at best no better than "common sense"; and (2) the devices whereby we may attempt to predict changes in behavior do not have the benefit of empirical validity, and efforts to validate our tests and our practices have detracted more from our confidence than they have contributed.

When confronted with such a suggestion of weakness, one seeks reassurance and one hopes that recent research trends provide evidence of strength. Accordingly, some of the current literature was reviewed.

Of the various journals that publish research reports which are relevant to the problems and interests of clinicians, the *Journal of Consulting Psychology,* the *Journal of Abnormal and Social Psychology,* and the *Journal of Clinical Psychology* are perhaps best known and most generally read by psychologists. The content of research published in these and other journals is the subject of frequent, well-known, and competent reviews. In our review, we were not concerned with *content;* we assumed that, for the most part, the content of the reports was relevant to the interests of clinical psychologists. Our review covered one year, 1952, and was designed to reveal the degree to which and the respects in which this research was explicitly or implicitly related to the problem of behavior change. An outline (see Table 1) was prepared which could provide a classification of research reports with respect to the descriptive function, the predictive function, and the change function of the clinical psychologist.

Parts of Table 1 may require explanation. In the *Description* part of the outline the various sources of descriptive information are classified. For example, "Observation" as a source of descriptions may be informal, or it may be formal and include the use of check lists and rating scales. The unique feature of observation is that the observer participates to a relatively small degree. In sources classified as "Interview" the data could be recorded in any form, but there was an interview relationship involved. In the "History" classification the information descriptive of the individual's past circumstances could come from any prior source. It could also come from any concurrent source other than concurrent observations, interviews, or tests. The "Tests" included any standard situation for sampling the subject's reaction. The essential characteristic of our test source was that the situation was standard, relevant to some specified reactions of the subject, and designed to provoke reactions from the subject.

TABLE 1 *A survey of current research for clinical psychologists*

Classification of Research Approaches	Frequency of Various Research Approaches in Articles Published during 1952 (*J. consult. Psychol.*, *J. abnorm. soc. Psychol.*, and *J. clin. Psychol.*)
I. Description	
A. Observation	1
1. Temporal trend inferred	
a. From past	1
b. Toward future	
2. Concurrent material interrelated	9
B. Interview	1
1. Temporal trend inferred	1
a. From past	
b. Toward future	
2. Concurrent material interrelated	3 *
C. History	
1. Temporal trend inferred	
a. From past	
b. Toward future	
2. Concurrent material interrelated	35 *
D. Tests (or other standard situations for eliciting the subject's reactions)	28
1. Temporal trend inferred	
a. From past	2
b. Toward future	1
2. Concurrent material interrelated	110
II. Prediction	
A. Stability of inferred characteristic (either stable traits or trends of change)	
1. Success	8
2. Failure	1
B. Susceptibility to change as a result of a specific experience (including a successful therapy)	
1. Success	12
2. Failure	5
III. Therapy—change	
A. Change or attempt at change described only (The agent may or may not be specified and its effectiveness may or may not be anticipated)	
1. Success	9
2. Failure	2
B. Change or lack of change explained *ex post facto*	2
Non-research	35
Total	266

* Studies involving some other concurrent descriptive category also.

TABLE 2 *Summary of research approaches*

Principal Approaches	Per Cent of Occurrence	
Descriptive	83	
Trend inferred		(3)
No trend inferred		(80)
Predictive (includes no purely descriptive studies)	11	
Stability tested		(4)
Susceptibility to change (including therapeutic) tested		(7)
Therapeutic—change (includes no studies of differential susceptibility to change)	6	
Total	100	

NOTE: In the descriptive approaches all the historical sources and three of the interview sources occurred in studies which also involved test sources of data. Aside from this, there is almost no overlap. This is possible because the predictive and the therapeutic studies necessarily involve description, and for such studies no tally was made for description. Another kind of possibly overlapping count was eliminated by excluding from the therapeutic category all studies which involved an explicit prediction of differential susceptibility to change.

Usually, when behavior is described by any of the foregoing procedures, it is implied that the content of the description comprises a stable characteristic of the individual or of his behavior and that any change in this characteristic either is an exception or is due to some change in circumstances. This static emphasis is not invariable, however. In practical clinical work clinicians frequently infer a trend from the available descriptive material. The trend they describe is usually historical, but sometimes it is conceived in such a way as to anticipate the future behavior of the individual. Recognition of such trends was sought because it could be taken as an indication that the research was relevant to an aspect of the behavior-change interests of clinical psychologists.

The interrelation of concurrent material in some cases involved descriptive information from two or more of the classified sources of information. Almost all these overlapping cases were studies which involved both "History" and "Tests" information.

The *Prediction* part of the outline refers to studies wherein descriptive information which was gathered at a prior time is related with descriptive information gathered at some subsequent time. The intervening time interval may cover a period of incarceration, a course of psychotherapy, or a laboratory experience. Prediction studies which examine the stability of a characteristic were in some instances concerned with a static trait. In other instances they were concerned with a growth or change process in the individual. Predictions concerning this susceptibility to change refer to all manner of reports wherein some prior characteristic of the individual, either naturally occurring or experimentally induced, was predicted to qualify the nature of the individual's response, i.e., change, in some other situation. The prior characteristics of the individuals could include such extremes as symptoms or specific instructions by an experimenter. The situation in which change was differentially predicted could be therapy or could be a learning situation. Obviously, predictive studies involve description, but predictive studies were not classified under the descriptive heading merely because of this necessary fact.

The *Therapy* or *Change* section does not include all the studies which have to do with change. Specifically, it excludes any studies having to do with change wherein the differential susceptibility of the individual to change was specifically predicted. If we except the susceptibility-to-change predictions, the therapy section of the outline refers to studies where an actual change occurred or where evidence of change was sought in a presumably change-inducing situation.

As would be expected, a number of the articles did not have a research content. Their content has not been classified. Since none of them included any data, they could not be expected to contribute any new facts.

Referring to Tables 1 and 2, one sees that our recent research has little concern with behavior change. Eighty per cent of the articles are not only independent of the topic of behavior change, but also they are not even suggestive of an interest in the development or the modification of human behavior. Four per cent of them appear to be explicitly concerned with the stability of behavior. Only 13% of the studies show a concern for describing or predicting change. Either clinical work is not much concerned with behavior change, or the research published by and for clinical psychologists is not much concerned with clinical work.

Since an examination of our current research in clinical psychology fails to offer us reassurance, it may be worth our while to discuss briefly some of the broad areas of interest among clinical psychologists. Perhaps such a discussion may help us to think further about the relevance of changes in behavior for the work and for the research of the clinical psychologists.

THE ABNORMAL

Most frequently clinical psychologists are concerned with deviations of behavior. Such deviations are usually called abnormal. Unfortunately, there are numerous ways in which abnormal behavior has been defined.

Some authorities have suggested that some sort of definition of abnormal based on a statistical frequency is sufficient. For the most part these people are inclined to believe that all human attributes exist in varying degrees and that some degree of each of these attributes may be found in all people. They would say that, in cases where some particular degree of these attributes exists, the persons concerned may arbitrarily be classified as abnormal with respect to these attributes. In practice, however, such a point of view is difficult to employ because many of the behavioral attributes which characterize people who are regarded as abnormal or in difficulty cannot be confidently seen as an extension of well-known and explicitly defined continua which characterize everyone. Another difficulty in the statistical-frequency point of view lies in the fact that the nature of relevant characteristics of individuals, as well as the frequency of any degree of such characteristics, seemingly varies considerably from cultural subgroup to cultural subgroup. The norm requirements of the statistical-frequency point of view become most discouraging when we realize that it is possible to conceive of a very large number of behavioral continua for any one subculture.

There is also the point of view that what is normal and what is abnormal may be specified with respect to some sort of ideal; i.e., normal is considered to be what is desirable. Upon reflection it seems possible that the sources of such ideals of desirability may be quite numerous and would include political and ethical considerations and could change after every election. Of course, there are medical ideals. These seem to give no sure footing, however, because even in the aspects of medicine which seemingly are solidly founded and well-known change occurs, so that the ideal behavior for eating, bathing, sleeping, exercising, etc., seems to change from time to time. Departures from ideal behavior even in the form of symptoms vary in the seriousness of their implications from area to area, depending upon the kinds of diseases that may be found to be endemic.

There are those who would claim that, although one sees in the abnormal individual the same sort of behavior as is seen in the normal individual, somehow or other in the abnormal individual this behavior has a different meaning. The laws by which it is combined or regulated are claimed to be different from the laws applicable to normal behavior. Presumably one can understand the essential abnormal nature of behavior only when one can see how the behavior in question is abnormally interrelated. This point of view, although superficially attractive, is rather disappointing when examined critically because it implies that one knows how behavior or behavioral manifestations are normally interrelated.

Concerning the absolute value of clinical judgments of the content of behavior, almost anyone who has attended numerous clinics may come away with the feeling that almost all human attributes and all human life circumstances have at one time or another been regarded as essential factors in the development of, or crucial evidence for the development of, a marked and tragic mental pathology. Nevertheless, study of individuals who are not and have never been patients indicates marked prevalence of the handicaps, disappointments, pressures, conflicts, uncertainties, abuses, and disillusionments which are claimed in the clinic as sufficient explanation for the development of pathological conditions.

Perhaps it is time to abandon the distinction between normal and abnormal; perhaps it is not truly relevant to the work of the clinical psychologist. To do this need not be construed as an abandonment of a goal of clinical psychologists. Obviously, it is not suggested that we should disregard the *differences in behavior* which we have attempted to relate to a concept of the abnormal. It is suggested instead that we give primary consideration to the differences in behavior from the particular standpoint of the conditions of their origin and change. We know that there are forms of behavior which are acutely distressing to the individual, and we know that also there are forms of behavior which may not be distressing to the individual but are acutely distressing to his associates. Perhaps the proper emphasis would be with the origin of these forms of behavior and with some concern as to why they cause distress. However, since there are so many ways in which an individual may be distressed or may distress others, it may be that *what* changes may not be practical. Perhaps our emphasis should be on *how* changes occur.

CONCEPTS OF PERSONALITY

The practical significance of description of current behavior states, especially abnormal behavior, is said to depend on the clinician's knowledge of personality. Such a point of view could be consistent with a behavior-change view of personality, but it obviously begs the question if personality is considered from the static-trait point of view.

The research study of personality may be arbitrarily classified into two main approaches. There are students of personality who approach their subject matter by investigating the manner in which various forms of behavior may be modified. The studies of interrelations between the appearance of forms of behavior and characteristics of the environment are often designed to give general understanding of the manner in which various forms of behavior may develop in the individual and the manner in which they may be changed. Unfortunately, when inferring modifiability of behavior many of these studies interrelate concurrent information descriptive of the individual with concurrent (or concurrently reported) information descriptive of the environment.

In a few instances the investigation of behavior modification has made good use of the two traditional forms for scientific investigation: i.e., the field or observational approach, where experiments of nature provide the data for drawing inferences or testing hypotheses; and the laboratory or analogical approach.

The other main approach in personality research has to do with the determination of traits. These traits are inferences or inventions of persons interested in certain aspects of human behavior. The inventions of these traits may be based on a study of the interrelationships among different kinds of observations, but the observations are usually concurrent and with few exceptions the traits are static. It is likely that the person inventing traits will succeed in sampling individuals and behavior which are relevant to his concern. He may invent traits on the basis of apparent relevance to political issues, personal values, or the way in which a person behaves in a social situation. All such inventions are arbitrary and serve only to provide aid for describing individuals in respects which may be of interest only to the persons who share the concern of the individual who invents the traits. It can scarcely be claimed that they have any basic significance that involves intrinsic, or ubiquitous, characteristics of human behavior.

ABILITY

Although the concept of personality is vague and is approached by many psychologists with doubt and inconsistency, it is customary among most psychologists to approach the concept of ability with self-assurance. This self-assurance may, however, be ill founded, and it s quite possible that many psychologists get an altogether unwar-anted feeling of confidence from the apparent objectivity and recision of the traditional statistical manipulations to which they ubject the kinds of performance records gathered by test procedures.

At present much of test development appears to be essentially tech-ological, more concerned with the technique of sampling behavior han with behavior *per se*. It would appear that the broad relevance nd general conditions for the behavior sampled by the test concern he test maker very little. He tends to express an almost exclusive nterest in the reliability of his test score and the correlational validity f the test score with reference to some static criterion. This interest, lthough technically necessary, is more statistical than psychological n nature. If tests were used only in the context in which they were levised and predictions were made only for groups and with respect o the criteria with which they are validated, it is quite probable that he use of tests could proceed without serious abuses. Unfortunately, owever, regardless of the manner in which the tests are made and heir original purpose, they are soon applied to situations different rom the original situation and to the evaluation of individuals. This atural effort to use the test in a general and flexible manner is ac-ompanied by a wide variety of abuses.

What determines an individual's relative goodness with respect to a est score may occasionally be the subject of some speculation, but with a few conspicuous exceptions it is rarely a subject of systematic exploration. The conditions of the individual's present and past under which a specified validity may be ascribed to the test are un-explored. It is obvious that many factors may be determining the relative test standing of an individual.

Even when current determiners of the test score have been recog-nized by either the maker of the test or the person who uses it, the manner in which a person's test score is likely to change as a result of subsequent experiences is unknown. It seems probable that indi-viduals vary greatly in their susceptibility to change, and this may be an important limiting factor in test validity as we now determine it. In college guidance work it is common to observe that a person's test score or a person's performance which the test presumes to anticipate may change greatly as a result of additional specific experience. Such changes can result from some insight that the individual acquires con-cerning his role in the situation, from new motivation, or from a more healthful affective reaction. So in many testing situations it is quite possible that an estimate as to whether or not the score is an optimal score or minimal score may be as important as the score itself.

Test makers seemingly are little concerned in this, and those who use the tests apparently are satisfied to be guided by informal guesses. Until we recognize that in general we do not know to what situations the test may be safely generalized and to which ones it may not be safely generalized our practical testing procedures can lead us into serious unprotected errors. Perhaps we are justified in saying only that our tests are based on a sampling procedure, and the degree to which we can predict and otherwise generalize is determined entirely by the resemblance between the situation under which we test (sample) behavior and the situations to which we wish to generalize.

THERAPY

The problem of change and the individual's susceptibility to change which has been suggested by our discussion of personality and ability may have a central relevance to clinical work, particularly since clinical work is concerned with therapy. Clinical psychologists in some situations conduct therapy themselves, in other situations they are expected to recommend appropriate kinds of therapy, and in still other situations they are expected to predict the outcome of possible therapies. All these responsibilities involve judging the individual's susceptibility to change as a result of certain experiences. This is ironic because one of the things about which we know the least is the conditions under which individuals change.

There are changes which can be brought about in the individual with varying degrees of success but without any clear rationale to account for success or lack of success. For example, in a therapeutic situation it is often most important to give the individual some knowledge of his environment: what he can anticipate tomorrow or the next day, how he can expect people to react to him, what he can expect to find in various situations, etc. Just how these changes are brought about in the individual is not known, but procedures for conveying such information seem to be our heritage and include pantomime, the spoken word, and the printed page.

Apparently knowledge concerning the individual himself can also be imparted to him, and this is often accompanied by desirable behavioral changes. Certain special preconditions must usually be established before self-knowledge is imparted or gained. Unfortunately, it seems that the therapist must decide upon the nature of these necessary preconditions by intuition and must evaluate them informally. After such evaluations of preconditions are made the question of whether the therapist conducts this teaching (or guides the learning) in a way which is more efficacious than any randomly selected way remains unanswered. As he conducts his therapy he has neither the benefit of an explicit rationale nor the support of empirical studies. Let us hasten to note that at the conclusion of any phase of therapy he may review his records and make an *ex post facto* application of some theoretical formulation to what has transpired, but can he use any current theory to plan and successfully anticipate the next phase of his therapeutic efforts?

CONCLUSION

Many of the things that we do or say in relation to the words "abnormal," "personality," "ability," and "therapy" lead us to both a practical and a conceptual emphasis on the role of behavior change. The concept of abnormal seems to have served us poorly, and we are tempted to turn to a point of view which does not emphasize a distinction between normal and abnormal, but which seeks to understand the manner in which behavior and its personal and social values emerge and are modified. Our consideration of personality leads us to skepticism concerning the value of descriptions which are designed to serve no explicit purpose; our hope for progress in the study of personality includes an understanding of the conditions necessary for constancy and for change in behavior. Our remarks concerning ability and its assessment center about success and failure in prediction; such a discussion of prediction must turn to a further consideration of change in human behavior and of the problem of individual differences in susceptibility to change. Finally, when we discuss therapy, we find that our description of therapeutic change must involve the conditions for necessary prior changes, and we suspect that any plan of therapy must include evaluations (either explicit or implicit) of the individual's susceptibility to diverse specified therapeutic conditions. Thus it appears as we emerge from our brief discussion of concepts and practices of clinicians that our emphasis on change is greater than it was when we began.

We become increasingly aware that the content of the behavior change which has our interest depends upon our values and our purposes, and they in turn we see to have but a regional and temporal significance. Accordingly, it becomes apparent that our conceptual, theoretical, and, if you please, scientific interest turns to the *how* of behavior change, and the content or the *what* of the change in behavior is but an accident of time and place. An emphasis on the *how* of behavior change is not only an implicit expression of a hope but also an implicit application of a premise which is familiar in much of the work of psychology; specifically, it implies that *change* in behavior, regardless of content, may be understood (i.e., may be explained, predicted, and produced) by an employment of a single set of concepts. This premise implies that *one* psychology of behavior change may be sufficient. The antithesis to this premise would be either that no theory of behavior change is possible or that a different theory of behavior change is necessary for different behavioral contents. It would seem that a denial of the premise that one theory is sufficient for behavior change is an admission of a possible chaos in the study of human behavior, and it implies that psychology can be no more than a technology for sampling responses and for generalizing from such a sample of responses to similar samples of responses.

If it is reasonable to think of the psychology of the clinician as the psychology of behavior change, we must acknowledge that we have been, and are, extraordinarily coy about studying behavior change. It is obvious that we have tarried long with descriptions of behavior states and have been obsessively concerned with the *what* or the *content* of samples of concurrent behavior. If description of particular contents has been our purpose, we not only have often failed to say why we wish to describe (we have neglected to say how we hoped our description might be useful), but it can scarcely be claimed that we have done well in our descriptive efforts. Even our techniques of description seem to be content bound, and if there are broadly general considerations which can be applied to any problem of behavioral description most of us would have difficulty in verbalizing these principles in any succinct and mutually acceptable manner.

It is obvious that if we are to study behavior change we must describe what is changing in order to have referents in our discussion of change per se. It seems, however, that the time has come when we must acknowledge the technological nature of our descriptive tasks and explicitly develop the principles of description *per se* and not be content to publish anecdotes (numerical or otherwise) of our adventures in describing behavior. One wonders if the time has not come also to suspect that description without concern for change is likely to be an evasion of our practical and scientific responsibility.

Occasionally, one hears clinicians and other students of human behavior say, "We must turn to learning theory." Although such statements imply an acknowledgment of the limited value of much of our descriptive research, they also imply that we have but to "turn" to learning theory to get our answers. As yet there is no reason for believing that learning theory has the answers, and it is possible that systematic and thoughtful observation of the changes in human behavior under diverse conditions can contribute more to learning theory than learning theory can at present contribute to the work of the clinical psychologist. Clinical psychologists should not underestimate their opportunities for the careful description of *behavior change* and correlation of such *behavior change* with prior characteristics of an individual and with current features of his environment. Such studies can be conducted as procedures for examining hypothetical deductions from some recognized theory, or they can be frankly exploratory and possibly the basis for inferences. It seems, however, that the study of changes (naturally occurring, educational, therapeutic, or laboratory) is not only the urgent responsibility but also the golden opportunity of those who do research in human behavior, particularly in clinical psychology.

Index