First published in *Psychological Review 65*(1), 1958, 22–33.[1]

"What is Learned?"—An Empirical Enigma

Although cognitive and stimulus-response interpretations of learning have traditionally been regarded as antipodes in the theoretical disputes of behavioral psychology, it has become increasingly apparent that there are few, if any, predictive *behavioral* differences between these positions. Superficially, one might expect gross disagreement between an S-R theory, which holds to a direct attachment of responses to stimuli, and a cognitive or "expectancy" theory which maintains that conditioned stimuli elicit sensory processes, or surrogates thereof, previously associated with other stimuli. Yet when one reflects that the concept of "mediation response" (Osgood, 1953) or "anticipatory goal response" (Hull, 1931) allows a stimulus to evoke a CR whose proprioceptive feedback provides a mechanism by which a stimulus may have conditioned sensory consequences, it is no longer surprising that no one has yet succeeded in cooking up a "critical" experiment to settle the controversy.

It is entirely too rash, however, to conclude with Kendler (1952) that speculations as to "what is learned" are empirically meaningless. To do so is to overlook the crucial point that the known—or supposed—facts of behavior have driven S-R theorists from their older models *without* mediation responses to a more complex theoretical analysis which is, within rather wide limits, *formally* identical with expectancy theory. There *are* important predictive differences between the old, unmediated, S-R theory on the one hand, and mediated S-R or expectancy theory on the other. Hence, if by "What is learned?" we are concerned not with the connotative imagery associated with theoretical terms but with the formal properties of learning theories, it would appear that this question does have empirical significance. The fact that behavior theories of widely divergent conceptual origins appear rapidly to be converging to a common formal structure strongly implies the existence of certain *empirical* principles of behavior which necessitate this partic-

[1]This article is based upon a dissertation submitted to the Faculty of the Division of the Biological Sciences, the University of Chicago, in partial fulfillment of the requirements for the Ph.D. degree. The writer wishes to express his gratitude for the generous support of the National Science Foundation, whose fellowship program has subsidized the writer's graduate studies.

Editorial Note: In Rozeboom (1997, p. 368, fn. 23) the author describes this article as his "personal epiphany" on the existence of a logic of discovery. He says: "What came as revelation to me was realization that although non-mentalistic S-R mediation mechanisms could in principle account for a certain prospective transfer-of-training phenomenon that commonsense would take to manifest a mentalistic "idea" mediating between external stimuli and behavior, this phenomenon would demand explanation by a certain structure of mediation with indifference to whether that structure has a mentalistic, S-R mechanistic, or some other embodiment. (See Rozeboom, 1970, pp. 120–122)".

ular structure. And if such is indeed the case, we are confronted with a problem: Exactly what *are* the empirical principles which give rise to mediational theories of learning?

The question just posed is not as simple as it might appear. Every graduate student of behavioristics is familiar with, e.g., the "latent learning" experiments which were supposedly critical for expectancy theories, and whose explanations in S-R terms draw so heavily on mediation responses. But an experiment is not an *empirical principle*; by the latter, we understand a lawful relation among observation variables which, in conjunction with other empirical principles, determines the outcome of an experiment. To understand the empirical bases for mediation hypotheses, it is insufficient to know what experimental outcomes favor a mediation hypothesis. We must also know the laws which these results exemplify.

Let me elaborate on this point, for it is an important one. There appear to be two main approaches through which one can seek understanding of natural phenomena. One—the theoretic—is to hypothesize about the underlying unobserved entities or processes and their relations which are casually responsible for the gross observed effects. The other—the empiristic—is to seek discovery of a set of dependable empirical covariations, in terms of which more complex phenomena may then be analyzed. There is no need here to weigh the relative merits of these approaches.[2] What is important to note is that there is no reason to doubt that any behavioral phenomenon can be analyzed as an instance of one or more purely empirical principles, even though these, in turn, may call for explanation in theoretical terms. The importance of such "brute facts" of behavior for empiricist and theorist alike is appreciated readily enough by reflection on the position in behavioral psychology of the empirical laws of conditioning, both classical and instrumental, the facts about deprivation and activity level, etc. Such principles may be regarded as what we *know* about behavior, in contrast to what we *surmise* (though, of course, in the last analysis this is less an absolute distinction than a matter of degree), and it is this hard core of *facts*—by which I mean not merely data, which are gathered readily enough, but the much less obvious empirical covariations which these exemplify—upon which the *science* of behavioristics is erected.

The question which I have raised is whether the problem, "What is learned?", is not reflected on the empirical level as a question about the laws of behavior. I have already argued that the structural convergence of alternate learning theories does, in fact, suggest this to be so. The question still remains, however, as to whether the particular relations responsible for mediation hypotheses are merely

[2] If space permitted, I would argue that theoretical concepts are—within sharply circumscribable limits—not merely justified, but are more or less *compelled* by certain kinds of empirical data.

special instances of laws already known, *or whether, perhaps, a previously unrecognized basic principle of behavior is involved.* I shall attempt to demonstrate that the latter is, indeed, the case—or more accurately, that the empirical basis for mediation hypotheses is a fundamental behavioral principle *if*, in fact, it exists. For, as will be shown, the problem at hand carries us into a new, essentially unexplored, dimension of research.

I shall not attempt to display the empirical problem of "What is learned?" through analysis of the traditional latent learning and reasoning experiments. To do so would be to become lost in a mass of obscuring details, and to suggest that the principle for which we are seeking manifests itself only in special circumstances. Rather, I shall demonstrate that *the empirical possibility which gives rise to mediation theories of learning is an intrinsic aspect of each and every instance of conditioning phenomena.* The known facts of conditioning are traditionally phrased in a way that conceals a highly critical ambiguity concerning what *are* the facts of conditioning. In the next section, I shall attempt to formulate the conditioning paradigm in complete generality, and to exhibit the correlated empirical mediation problem, without whose answer a resume of the facts of conditioning is as incomplete as a specification of geographical location which gives latitude but not longitude. But to illustrate my point and to illuminate the general analysis, I will first cite a specific example.

Perhaps the oldest empirical principle of behavioristics is that of classical conditioning. This may be expressed:

> (I) *If an organism repeatedly encounters, under suitable circumstances, a stimulus,* $\mathrm{S_c}$, *immediately preceding a second stimulus,* $\mathrm{S_u}$, *which evokes response* R, *then* $\mathrm{S_c}$ *will come to evoke* R.

With, some qualifications (chiefly in regard to the similarity of CR and UR), (I) may stand as a close paraphrase of conventional formulations of classical conditioning, and *seems* to assert no more than is justified by generalization from known data. But compare it with the following formulation:

> (II) *If an organism repeatedly encounters, under suitable circumstances, a stimulus,* $\mathrm{S_c}$, *immediately preceding a second stimulus,* $\mathrm{S_u}$, *then* $\mathrm{S_c}$ *will come to evoke the same response(s) as elicited by* $\mathrm{S_u}$.

Both (I) and (II) are equally supported by the facts of conditioning, so far as these are known; *yet* (I) *implies that classical conditioning establishes a bond between* $\mathrm{S_c}$ *and* R *which is independent of the subsequent behavioral effects of* $\mathrm{S_u}$, *whereas* (II) *implies that any modifications in the response-evocation properties of* $\mathrm{S_u}$ *will be passed along to* $\mathrm{S_c}$*—i.e., that classical conditioning establishes a functional equivalence between conditioned and unconditioned stimulus.* Since the behavioral effects of the US have almost invariably been left unaltered in conditioning experiments, there is absolutely no justification for assuming (I) to be any more representative of the empirical facts than (II). Yet it is traditionally assumed by empiricists and S-R theorists that the conditioned reflex is a "direct" or "mechanical" attachment of response to stimulus—i.e., that (I) is the case. The truth is that we *simply don't know* what the relationship is between the responses to the CS and US subsequent to the conditioning operation, despite the fact that the value of this information, both for theoretical and applied purposes, can scarcely be overestimated.

The moral that I want to draw is that our knowledge of the empirical facts of classical conditioning is incomplete—not in the relatively trivial sense that we are not yet acquainted with all the parameters of its occurrence, but in the extremely basic sense that we have only partial knowledge of its empirical *consequences*. Moreover, unless we explicitly stipulate that the behavioral effects of the US remain constant, it appears extremely difficult to express what we do know about conditioning without inadvertently prejudging the results of experiments yet to be conducted. (Such prejudgments undoubtedly account for our traditional assumption that a conditioned reflex, once established, is independent of the subsequent effects of the US.) But the relation between the response-evocation properties of the CS and those of the US constitutes an empirical principle yet to be determined, and it is the likelihood that a significant relationship does in fact exist, along with others of its kind, that has encouraged the development of mediational theories of learning.

THE GENERALIZED CONDITIONING PARADIGM

The preceding example demonstrates how, for classical conditioning, the problem, "What is learned?", is not primarily a theoretical issue at all, but a straightforward query about the brute facts of behavior. As we shall now see, a similar problem exists for every form of conditioning. To show this in complete generality, it is first expedient to assimilate the various instances of conditioning—conditioned reflex, instrumental conditioning, conditioned reinforcement, etc.—under a single rubric.

Most, if not all, of those instances of learning to which the term "conditioning" has been applied, as well as foreseeable extensions of this term, may be subsumed under the formula of a stimulus, S_c, acquiring certain behavioral effects (for a given organism) through participation of S_c in certain contingency relations with other environmental events. By "behavioral effect" I mean any change, transient or enduring, observed or inferred, in the behavioral attributes of an organism which results from reception of the stimulus by that organism. The simplest and best known behavioral effect of a stimulus is, of course, response evocation, but more complex consequences of stimuli for behavior are also known. One such effect is (instrumental) "reinforcement," which confers upon a stimulus the ability to produce instrumental conditioning. Still another is drive arousal or reduction, which differs from simple response evocation in that the former is characterized by a change in activity level which may be expressed in a variety of behaviors. There is evidence that these effects may be acquired and lost in manners vary similar to the acquisition and loss of response evocation.

Since environmental events are presented to the organism in terms of stimuli,[3] we may formulate the conditioning schema as follows: *A phenomenon is known as "conditioning" when it may be described as the acquisition by a stimulus,* $\mathrm{S_c}$*, of a behavioral effect,* $\mathrm{E_c}$*, through participation of* $\mathrm{S_c}$ *in a contingency relation,* $\mathrm{T(S_c, S_u)}$*, with an "unconditioned" stimulus,* $\mathrm{S_u}$*, whose behavioral effects are* $\mathrm{E_u}$*.* (Although the origins of "the term "unconditioned stimulus" lie in the innate autonomic reflexes of Pavlovian conditioning, it is terminologically convenient to regard the stimulus through which conditioning is accomplished as an unconditioned stimulus, even though it may have received its own behavioral effects through a previous conditioning procedure.[4]) Let "E is a behavioral effect of stimulus S" be abbreviated "$S \to E$." Then if $S_u \to E_u$ and $T(S_c, S_u)$ together causally imply $S_c \to E_c$, $T(S_c, S_u)$ is known as a "conditioning operation," while if $S_c \to E_c$ is due to a conditioning operation, S_c and E_c are a "conditioned stimulus" and a "conditioned effect," respectively. The role of E_u in the conditioning procedure seems to be that of determining the nature of the conditioned effect, E_c, since the conditioning powers of a given S_u are altered as its E_u is altered.

It must clearly be appreciated that the generalized conditioning paradigm, so formulated, in no way asserts a *law*; it merely describes those generic empirical features which are exemplified by conditioning phenomena. So far, nothing has been said about the kinds of behavioral effects involved, nor have we specified any details of the conditioning operation, $T(S_c, S_u)$. There will be as many species of conditioning as exist behavioral effects and contingency relations with the appropriate consequences. Neither, as yet, has the relation between the unconditioned effect, E_u, and the conditioned effect, E_c, been specified; we shall turn to a detailed analysis of this following a brief summary of the two major known types of conditioning operations.

*Classical, Pavlovian, or type-*S *conditioning.* In this instance, the occurrence of the unconditioned stimulus, S_u, is made contingent only upon the occurrence of the conditioned stimulus, S_c. So far as is known, classical conditioning can be accom-

---

[3]Since *behavioral* manipulations of the organism consist of *environmental* modifications, the independent variables of behavioristics, including drive operations, can always be formulated in terms of stimulus situations, although certain methodological complexities arise in the case of the deprivations.

[4]It would be preferable to call S_u the "conditioning" stimulus, but this makes for excessive confusion with "conditioned" stimulus.

plished whether the behavioral effect of S_u is innate or itself conditioned, though for technical reasons, higher order classical conditioning is difficult to demonstrate empirically (Skinner, 1938, p. 244f.), and is, perhaps, but weakly established (Razran, 1955). With the behavioral effects of S_u held constant, the conditioned effect, E_c, appears to be similar to the unconditioned effect, E_u, when E_u is response evocation, positive or negative reinforcement, or drive evocation (e.g., Calvin, Bicknell, & Sperling, 1953b, Danziger, 1951, Hilgard & Marquis, 1940, Keller & Schoenfeld, 1950, Miller, 1951, and Skinner, 1938). Available evidence (e.g., Calvin, Bicknell, & Sperling, 1953a and Miles & Wickens, 1953) tends to weigh against the conditioned acquisition of drive reduction.

*Instrumental, or type-*R *conditioning.* When the occurrence or cessation of a stimulus, S_u, is made contingent upon emission of a response, R, in the presence of a stimulus, S_c—i.e., the occurrence or cessation of S_u is dependent on the joint occurrence of S_c and R—the subsequent likelihood of R as a response in the presence of S_c may be altered, depending upon the reinforcement value of S_u (e.g., Hilgard & Marquis, 1940, Keller & Schoenfeld, 1950 and Skinner, 1938). When the onset of S_u contingent upon R in the presence of S_c results in the increment of R as a response tendency to S_c, then S_u is known as a "positive" reinforcer. When it is the termination of S_u that strengthens $S_c \to R$, then S_u is termed a "negative" reinforcer. The onset of a negative reinforcer contingent upon R in the presence of S_c is known to decrease the strength of $S_c \to R$, perhaps through instrumental strengthening of competing responses (e.g., Dinsmoor, 1954). During instrumental conditioning, the reinforcement value and probably other behavioral effects (other than drive reduction) of the reinforcing stimulus also apparently tend to become conditioned to S_c, undoubtedly through a type-S process. Thus R is only a portion of the total effects, E_c, acquired by S_c during instrumental conditioning. (Note, incidentally, that classical conditioning may be conceived as a special, or degenerate, case of instrumental conditioning in which the instrumental response, R, is the null-response—i.e., S_u is contingent upon S_c plus no particular response.) There is still much that is unknown about the parameters of instrumental conditioning, particularly in regard to the participation of drive.

THE PROBLEM OF CONDITIONED GENERALIZATION

In predicting the behavioral effects of a conditioned stimulus, it is insufficient to know merely the stimulus contingencies in which S_c has participated; we also need to know the behavioral effects of S_u. That is, E_c is a function not merely of $T(S_c, S_u)$, but also of E_u. But what do we know of this function empirically?

The extent of our knowledge here has already been summarized (perhaps overtersely) above. We know that *so long as the behavioral effects of* $\mathrm{S_u}$ *remain con-*

stant, roughly $E_c = E_u$ for classical conditioning,[5] and roughly $E_c = R + E_u$ for instrumental conditioning, where R is the instrumental response (or, in certain instances, its inhibition). But this leaves unspecified *which* behavioral effects of S_u determine E_c, for, in principle, the effects of S_u can vary with time. At a given moment, are the conditioned effects of S_c determined by the effects of S_u *at that same moment*, or by the effects of S_u *at the time of conditioning*? (There are, of course, still other possibilities.) Either alternative has as much a priori likelihood as the other, and only data from experiments in which the effects of S_u are altered subsequent to conditioning will permit decision between them.

The analysis so far has been conducted in a quasi-formalistic manner in an effort to exhibit the problem of the functional dependence of. conditioned effect upon unconditioned effect as, logically, an intrinsic aspect of conditioning phenomena. Let me now put it straightforwardly. *When a stimulus,* $\mathrm{S_c}$*, is conditioned to a behavioral effect,* $\mathrm{E_c}$ *through contingencies of* $\mathrm{S_c}$ *with an unconditioned stimulus,* $\mathrm{S_u}$*, whose behavioral effect is* $\mathrm{E_u}$*, what is the functional dependence, if any, of* $\mathrm{E_c}$ *upon* $\mathrm{E_u}$ *after conditioning?* For greater clarity, we may formulate the problem as an experimental paradigm.

Phase 1. Condition a behavioral effect, E_c, to a neutral stimulus, S_c, by making stimulus S_u, with behavioral effect E_u, contingent upon S_c alone, or upon a response, R, in the presence of S_c.

Phase 2. Extinguish E_u as an effect of S_u and condition to S_u a new effect, E'_u. Extinction of E_u is not essential, since the objective here is only to modify the effects of S_u. However, extinction of E_u, when this is possible, tends to maximize the difference in the effects of S_u before and after modification, and hence to maximize evaluation of the results in Phase 3.

Phase 3. Test S_c for its influences on behavior.

There are three major alternative results which might be anticipated as the outcome in Phase 3. These are not logically exhaustive, but would seem to cover the intuitive expectations of the behaviorist.

A. Stimulus S_c may continue to have E_c as its behavioral effect, and have gained no new effect. This indicates that under the parameters of the experiment, the behavioral effects of a CS are independent of the subsequent effects of the US.

[5]Cognitive theorists have shown an inordinate concern over the apparent fact that the classically conditioned response tends to differ somewhat from the unconditioned response. There is no more a priori reason why E_c should be identical with E_u than some other fixed function of E_u at the time of conditioning—neither is any less "mechanical" than the other. The difference between E_c and E_u lends no more special credence to expectancy theories than does the even greater lack of similarity between E_c and E_u during instrumental conditioning. What *is* important for a mediation theory is whether the behavioral effect of S_c, *whatever* this may be when E_u is held constant, shows any tendency to covary with E_u after conditioning.

Such a result is commensurate with a nonmediational theory of learning.

B. S_c may have lost E_c as an effect, but instead, have acquired effect E_c', where E_c' is that behavioral effect which S_c would have acquired during the initial conditioning had E_u', rather than E_u, been the effect of S_u at that time. This would imply that, under the parameters of the experiment, the behavioral effects of a conditioned stimulus continue to be determined by those of the stimulus through which conditioning was achieved, so that, empirically speaking, the conditioning operation has established a stimulus-stimulus relationship in that the organism's responses to S_c take account, so to speak, of the current behavioral significance of S_u. Theoretically, this would imply that the relation between a CS and its effect is mediated by a process whose behavioral properties are under control of the US.

C. S_c may retain effect E_c, perhaps more or less attenuated, and also acquire some degree of E_c', thus implying that the effects of S_c are only partly dependent upon those of S_u under these parameters. The correlated theoretical interpretation is that the behavioral effects of the CS are multidetermined, some involving a mediation process controlled by the US, and others being independent of it.

For convenient reference, it is desirable to find a distinctive title for the experimental paradigm just described. Since Outcomes B and C display a dependence of the behavioral effects of one stimulus upon those of another as the result of a conditioning procedure, the name "conditioned effect dependency" provides a literal description but is unpleasantly cumbersome. However, when the effects E_c and E_c' of the conditioned stimulus are similar, respectively, to those, E_u and E_u', of the unconditioned stimulus, the much simpler term, "conditioned generalization," is also appropriate. For the sake of descriptive expediency, therefore, let me stretch a point and refer to the generic design as the *conditioned generalization* paradigm, even though the effects of the CS and US need not literally be the same.

There are several important observations to be made here. First, let me emphasize once again that although the conditioned generalization paradigm is distinctive as an experimental procedure, the empirical question which it is designed to answer is *not* a problem separate from or secondary to the facts of conditioning. Quite the contrary, it is a question *about* the facts of conditioning. As was illustrated earlier for the conditioned reflex, the empirical laws of conditioning are traditionally formulated in a way that negatively prejudges the existence of conditioned generalization, with the result that any experiment which appears to exemplify this effect is regarded as a special and complex instance to be explained away, if possible, in terms of more "fundamental" empirical principles. But from the available data, there is no more reason to think that the primary empirical consequence of the conditioning operation is establishment of a rigid $S_c \rightarrow E_c$ connection than it is the formation of a generalization from S_u to S_c. We simply do not yet know what the facts are.

Secondly, it is important to realize that the question of conditioned generalization is not a unitary problem to be answered by a single experiment, but a correlate to *every instance* of conditioning. There is no particular reason to think that the degree of conditioned generalization must be the same for every type of conditioning. In fact, judging from what data there are, it seems probable that the degree of conditioned generalization is not constant even for a given type of conditioning, but is highly dependent upon the particular parameters of the learning situation.

The next point is an extension of the first two. Since every type of conditioning is a conditioned generalization phenomenon (if only in a trivial sense that E_c is a constant function of E_u), the problem of conditioned generalization is quite literally a new *dimension* in the empirical study of learning, rather than merely an unexplored sector of an old framework. As such, the empirical principles which will emerge from the experimental study of this problem are bound to have powerful repercussions for our understanding of behavior. Not merely are the empirical facts of readily foreseeable applied value; data concerning the dependence of conditioned generalization upon parameters yet to be discovered will also provide a new set of observation variables intimately correlated with intra-organismic processes for which at present we have only tenuous theoretical speculations. It is perhaps not unfair to claim that conditioned generalization is one of the most important problems now confronting behavioristics.

Finally, it should be apparent that it is the facts of conditioned generalization, whatever they may be, which solve the old enigma, "What is learned?"[6] Far from being a theoretical blind alley, this is a straightforward question about the empirical consequences of the conditioning operation.

THE EVIDENTIAL BASIS FOR CONDITIONED GENERALIZATION

I have contended that it is belief in a principle of conditioned generalization that has motivated development of mediation devices such as expectancies and fractional anticipatory goal responses. Since rigorous demonstration of the dependency of theoretical mediation elements upon this kind of empirical covariation calls for a technical and fairly extensive logical analysis, I shall instead substantiate this

[6]Some writers (e.g., MacCorquodale & Meehl, 1954) have interpreted the question, "What is learned?", as concerning whether the conditioned response is a motor discharge or an achievement. However, this would seem to be more a generic problem in regard to the nature of the response than about learning as such. Since, as McCorquodale and Meehl point out (p. 219), S-R theorists are no more committed to, and in the main have no more subscribed to, a motor discharge conceptualization of the response than have cognitivists, the *dispute* over what is *learned* must lie in something other than this.

claim by showing that a large proportion of those experiments traditionally cited as critical for mediation theories are instances of the conditioned generalization paradigm. These may be grouped into three major categories: (*a*) "secondary generalization," (*b*) "preconditioning," and (*c*) a more loosely organized set of experiments which may be called "secondary instrumental conditioning."

Secondary generalization. In 1933, Shipley (1933) reported that if, for human subjects, an eyeblink is conditioned to a flash of light with a tap on the cheek as the US, and a finger flexion is subsequently conditioned to the tap (shock as the US), the light flash will be found to evoke finger flexion. On the basis of this single experiment, Hull (1934, 1943) proclaimed the principle of "secondary generalization," which hypothesizes that if two stimuli, S_1 and S_2, evoke the same response, R, then another response, R', conditioned to one of the stimuli, say S_1, will also be evoked by the other, S_2. This formulation differs somewhat from that of conditioned generalization in that, for secondary generalization, S_1 and S_2 have not necessarily had to participate in a common contingency relation. However, the case of secondary generalization, wherein the stimulus S_1, to which R' is subsequently conditioned, has been used previously as the US for conditioning R to S_2, as is true in the Shipley experiment,[7] is also, obviously, an instance of the conditioned generalization paradigm.

Shipley's experiment and results have been successfully repeated by Lumsdaine (Hilgard & Marquis, 1940, p. 230), and a similar design by Graham (1944), with dogs, also showed positive results. Thus the evidence is rather convincing that at least *some* degree of conditioned generalization is sometimes established during classical conditioning, although Razran (1955, p. 328) cites certain Russian studies that cast doubt on its ubiquity. These experiments do not permit evaluation of the strength of the effect, however; in particular, they do not permit decision between alternatives B and C of the conditioned generalization paradigm, nor whether conditioned generalization is the rule, rather than the exception, for classical conditioning. It should be added that the Shipley-Lumsdaine-Graham experiments appear to be the only available evidence directly relevant to the secondary generalization hypothesis, despite the important role this supposedly empirical principle has played in mediated S-R theory.[8] The notion of secondary generalization has

[7]Hull (1934), (1943, p. 192 f.) discusses Shipley's 1933 data as though Shipley had not conditioned finger flexion to tap on cheek, and as though transfer of finger flexion to light flash were due to shock evoking both eyeblink and finger flexion. Such an interpretation would be highly favorable to the existence of secondary generalization and the response theory of mediation, but it is definitely *not* in accord with the facts of the experiment.

[8]A later experiment by Shipley (1935), in which S_1 and S_2 were separately conditioned to R, appears to support the existence of secondary generalization as distinct from conditioned generalization. However, close analysis of the control data in this and the 1933 experiment reveals the 1935 results to be ambiguous. The preconditioning experiment by Wickens and Briggs (1951), mentioned later, might be interpreted as an instance of secondary generalization, but only if we

been occasionally invoked in other experimental contexts, such as generalization gradients among symbolically related stimuli, but more as an explanatory device than an object of experimental demonstration. Thus there are no grounds for believing in a principle of secondary generalization as distinct from conditioned generalization.

Preconditioning. The preconditioning paradigm begins with repeated pairings of two supposedly neutral stimuli, S_1 and S_2, and subsequently conditions a response, R, to one of them, say S_1. Preconditioning is shown if S_2 now also tends to evoke R. The preconditioning design is a special case of the conditioned generalization paradigm in that the behavioral effect of S_1 during the pairings of S_1 and S_2 is the null-effect, or, more likely, an unidentified effect. This phenomenon was demonstrated fairly conclusively by Brogden (1939). Since then, a number of preconditioning experiments have appeared in the literature (Bitterman, Reed, & Kubala, 1953, Brogden, 1942, 1947, Chernikoff & Brogden, 1949, Karn, 1947, Silver & Meyer, 1954 and Wickens & Briggs, 1951), all but one (Brogden, 1942) reporting positive findings. But, the claims of Bitterman et al. (1953) notwithstanding, any attempt to evaluate the magnitude of the effect strongly suggests that the phenomenon is a minor one, at least judging from the published studies. Thus the empirical importance of preconditioning is still problematic.

Secondary instrumental conditioning. In this category, we may include those experiments which, following establishment of an instrumental response, modify the behavioral effects of the reinforcing stimulus. More specifically, the problem may be phrased: If an instrumental response is conditioned to a stimulus, S_c, through reinforcement by a stimulus, S_u, and the reinforcement value of S_u is now altered, will the response to S_c remain unaffected, or will it have adapted in a manner commensurate with the new behavioral significance of S_u? When S_u is positively reinforcing during conditioning, and is subsequently extinguished of positive value or aversively reconditioned, this design may also be called "secondary extinction" (Rozeboom, 1957), "latent extinction" (Seward & Levy, 1949), or "nonresponse-extinction" (Deese, 1951). A typical secondary extinction experiment conditions, say, a position habit in a T maze, one goal box of which contains food or water. The animal is then placed for a while directly in the positive goal box, which is now empty of the primary reinforcer, and where the animal may also be shocked. The experimental question is whether the position habit will appear weakened on the next trial. The habit may be analyzed as a response chain held together by secondary reinforcement deriving from the sequence of discriminative cues. Reconditioning directly in the goal box tends to extinguish or aversively recondition the terminal reinforcers of the sequence, and the next maze trial reveals

---

are willing to regard human behavior in which a stimulus elicits a response as the result of verbal instructions as an instance of "response evocation" in the same sense that we apply this phrase to infra-human behavior.

the immediate effect on earlier links of the response chain.

An experiment of this sort was first proposed by Tolman (1933), who obtained negative results. But shortly thereafter, Miller (1935) reported that shocking rats placed directly in a distinctive goal box inhibited a running response which had previously terminated in that goal box. Secondary extinction research subsequently languished until the appearance of a number of recent studies (Deese, 1951, Moltz, 1955, Rozeboom, 1957, Scharlock, 1954, and Seward & Levy, 1949, most of which, unfortunately, have technical flaws (cf. Moltz, 1955, and Rozeboom, 1957) which vitiate their interpretive significance. However, the evidence seems fairly definite that, under at least some parameters of conditioning, a lever-pressing habit in rats is quite independent of the subsequent value of the reinforcers through which the habit was established (Rozeboom, 1957).

An important variety of traditional latent learning research (*type* 5 [MacCorquodale & Meehl, 1954, p. 211]) may also be reduced to the secondary-instrumental-conditioning rubric. In this instance, the organism is given experience with a maze, with eliminable reinforcers either absent, or equally distributed to all end boxes. The animal is now introduced directly into one end box and given a previously unencountered reinforcement, either positive or negative. The problem is whether or not the animal's maze behavior will have adapted to the changed conditions on the next trial. This differs from more obvious cases of secondary instrumental conditioning only in that the role in establishment of initial maze habits played by the stimuli whose secondary reinforcement values are subsequently modified is not so clear. When no known reinforcers are obviously influential in the initial maze behavior, this design may be described as "instrumental preconditioning," since it involves a stimulus, S_u (say, an end box), as contingent on a response, R, in the presence of another stimulus, S_c (the antecedent portion of the maze), but where S_u has no readily identifiable reinforcement value prior to the reconditioning operation. A number of experiments of this sort have been reported (see MacCorquodale & Meehl, 1954, p. 211, for summary; also, Honzik & Tolman, 1936, Minturn, 1954, and Strain, 1953), with outcomes about equally divided between positive and negative. Thus again, the evidence appears to favor the existence of conditioned generalization under some circumstances, but more explicit evaluation of the empirical significance of these experiments is precluded by the difficulties in giving an unambiguous account of the initial maze learning in terms of conditioning principles.

In summary of the available evidence concerning conditioned generalization, it would thus appear that while the quantity of relevant data is very meager, they distinctly imply that parameters of conditioning exist under which at least *some* degree of correlation between the behavioral effects of CS and US is established. Unfortunately, the existent experiments happen, for the most part, to be of special designs which either employ a minimal or unknown unconditioned effect, E_u, or do not compare the relative strengths of E_c and E'_c subsequent to reconditioning of S_u. It is only when E_u and E'_u are strong and well defined, and the alternatives E_c and E'_c are distinct and quantifiable, that we can decide whether the *primary* result of a conditioning operation is establishment of a full-fledged generalization from US to CS, or whether the prevalent notion that conditioning attaches a specific effect to the CS is basically correct, and apparent cases of conditioned generalization are secondary phenomena to be explained as special instances of more fundamental empirical principles. It is earnestly to be hoped that the present analysis, by stripping from the empirical question of "What is learned?" its obscuring theoretical superstructure, and developing a general experimental paradigm which isolates the essential features of the problem, will stimulate research intended not so much to test a theoretical prediction as simply to learn the facts.

SUMMARY

I have contended that, far from being a blind alley, the question "What is learned?" is an extremely important empirical problem about the consequences of conditioning, and one which has been concealed by improper formulations of the known facts. We have traditionally assumed that a conditioning operation attaches a fixed behavioral effect, E_c, to the conditioned stimulus, S_c, whereas in fact, the conclusion that S_c will give rise to E_c subsequent to conditioning *is empirically justified only so long as the behavioral effects of the unconditioned stimulus,* S_u*, remain constant*. The question then arises as to the extent to which, subsequent to the conditioning operation, the effects of S_c are a function of those of S_u. I have elected to identify this as the problem of "conditioned generalization," since, in one of its forms, it asks whether or not a conditioning operation establishes a generalization from US to CS.

Since the empirical consequences of the conditioning operation constitute a problem extending over the entire range of conditioning phenomena, and one which has remained unexplored in the overwhelming majority of conditioning experiments, conditioned generalization defines a new dimension of behavioral research. What few data are available suggest that conditioned generalization does exist as a phenomenon, but are totally inadequate as a basis for judging whether it is a basic behavioral principle or merely a secondary effect. Because of the practical value of this information, and its crucial significance for the development of behavior theory, it is to be hoped that the conditioned generalization problem will soon be subjected to a concerted experimental attack.

References

Bitterman, M. E., Reed, P. C., & Kubala, A. L. (1953). The strength of sensory preconditioning. *Journal of Experimental Psychology*, *46*, 178–182.

Brogden, W. J. (1939). Sensory preconditioning. *Journal of Experimental Psychology*, *25*, 323–332.

Brogden, W. J. (1942). Tests of sensory preconditioning with human subjects. *Journal of Experimental Psychology*, *31*, 505–517.

Brogden, W. J. (1947). Sensory preconditioning of human subjects. *Journal of Experimental Psychology*, *37*, 527–539.

Calvin, J. S., Bicknell, E. A., & Sperling, D. (1953a). Effect of a secondary reinforcer on consummatory behavior. *Journal of Comparative and Physiological Psychology*, *46*, 173–175.

Calvin, J. S., Bicknell, E. A., & Sperling, D. (1953b). Establishment of a conditioned drive based on the hunger drive. *Journal of Comparative and Physiological Psychology*, *46*, 176–179.

Chernikoff, R., & Brogden, W. J. (1949). The effect of instructions upon sensory preconditioning. *Journal of Experimental Psychology*, *39*, 200–207.

Danziger, K. (1951). The operation of an acquired drive in satiated rats. *Quarterly Journal of Experimental Psychology*, *3*, 119–132.

Deese, J. (1951). The extinction of a response without performance of the choice response. *Journal of Comparative and Physiological Psychology*, *44*, 362–366.

Dinsmoor, J. A. (1954). I. The avoidance hypothesis. *Psychological Review*, *61*, 34–46.

Graham, D. T. (1944). Experimental transfer of conditioning in dogs. *Journal of Experimental Psychology*, *34*, 486–493.

Hilgard, E. R., & Marquis, D. G. (1940). *Conditioning and learning.* New York: Appleton-Century-Crofts.

Honzik, C. H., & Tolman, E. C. (1936). The perception of spatial relations by the rat: a type of response not easily explained by conditioning. *Journal of Comparative Psychology*, *22*, 287-318.

Hull, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review*, *38*, 487–506.

Hull, C. L. (1934). The concept of the habit-family hierarchy and maze learning. *Psychological Review*, *41*, 33–54, 134–152.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century.

Karn, H. W. (1947). Sensory preconditioning and incidental learning in human subjects. *Journal of Experimental Psychology*, *37*, 540–544.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of Psychology.* New York: Appleton-Century-Crofts.

Kendler, H. H. (1952). “What is learned?”—a theoretical blind alley. *Psychological Review*, *59*, 269–277.

MacCorquodale, K., & Meehl, P. E. (1954). Modern learning theory. In W. K. Estes et al (Ed.), *E. C. Tolman.* New York: Appleton-Century.

Miles, R. C., & Wickens, D. D. (1953). Effect of a secondary reinforcer on the primary hunger drive. *Journal of Comparative and Physiological Psychology*, *46*, 77–79.

Miller, N. E. (1935). A reply to “sign-gestalt or conditioned reflex?”. *Psychological Review*, *42*, 280–292.

Miller, N. E. (1951). Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley.

Minturn, L. (1954). A test for sign-gestalt expectancies under conditions of negative motivation. *Journal of Experimental Psychology*, *48*, 98–100.

Moltz, H. (1955). Latent extinction and the reduction of secondary reward value. *Journal of Experimental Psychology*, *48*, 98–100.

Osgood, C. E. (1953). *Method and theory in experimental psychology.* Oxford: Oxford University Press.

Razran, G. (1955). A note on second-order conditioning-and secondary reinforcement. *Psychological Review*, *49*, 327–332.

Rozeboom, W. W. (1957). Secondary extinction of lever pressing behavior in the albino rat. *Journal of Experimental Psychology*, *54*, 280–287.

Rozeboom, W. W. (1970). The art of metascience, or, What should a psychological theory be? In J. R. Royce (Ed.), *Toward unification in psychology.* Toronto: Toronto University Press.

Rozeboom, W. W. (1997). Good science is abductive, not hypothetico-deductive. In L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* New Jersey: Erlbaum.

Scharlock, D. P. (1954). The effects of a pre-extinction procedure on the extinction of a place and response performance in a t-maze. *Journal of Experimental Psychology*, *48*, 31–36.

Seward, J. P., & Levy, N. (1949). Sign learning as a factor in extinction. *Journal of Experimental Psychology*, *39*, 660–668.

Shipley, W. C. (1933). An apparent transfer of conditioning. *Journal of Genetic Psychology*, *8*, 382–391.

Shipley, W. C. (1935). Indirect conditioning. *Journal of Genetic. Psychology*, *12*, 571–591.

Silver, C. A., & Meyer, D. R. (1954). Temporal factors in sensory preconditioning. *Journal of Comparative and Physiological Psychology*, *47*, 57–59.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Strain, E. R. (1953). Establishment of an avoidence gradient under latent-learning conditions. *Journal of Experimental Psychology*, *46*, 391–399.

Tolman, E. C. (1933). Sign-gestalt or conditioned reflex? *Psychological Review*, *40*, 246–255.

Wickens, D. D., & Briggs, G. E. (1951). Mediated stimulus generalization as a factor in sensory preconditioning. *Journal of Experimental Psychology*, *42*, 197–200.