PSYCHOMETRIKA—VOL. 8, NO. 2
JUNE, 1943

A GENERAL THEORY OF LEARNING AND CONDITIONING: PART II

WALTER PITTS
THE UNIVERSITY OF CHICAGO

The second of two parts of this article extends a mathematical theory of non-symbolic learning and conditioning to cases where reward and punishment are involved. The preceding results are generalized to the case where stimuli and responses are related psychophysically, thus constituting a theory of transfer, generalization, and discrimination.

4. Instrumental conditioning and the effect of reward and punishment.

In cases of what is often called "instrumental conditioning," in which the connection between stimulus and response is impressed at least partially by a reward following the evocation of the correct response, or is inhibited by a punishment so placed, a variety of observational considerations would appear to indicate that suitably placed affective stimuli influence the reaction-tendency E_j directly, in the manner of true conditioning, rather than merely the threshold R_j; thus, for example, inhibition of a conditioning consequent upon postliminary painful stimuli seems rather to have the nature of true counter-conditioning than inhibition, both in relatively slow temporal coarctation, and stability under influences which generally effect at least temporary loss of most inhibitions. Conversely, reinforcement by subsequent reward appears to have most of the properties of ordinary conditioning. This has been evidenced in several studies, especially those of Youtz (55, 56),* and also Brogden, Lippman and Culler (3), Brogden and Culler (2), Bugelski (5), and Skinner (46), with a summary in (18). This being the case, when affective stimulation is involved in the experimental routine, we must modify the expression (10) to accommodate it.

Our procedure for deriving the effect of this factor will again be much like that of the previous sections. Considering the situation of

* The numbered references in the present paper are all to the bibliography in the previous section of the present discussion, *A General Theory of Learning and Conditioning, Part I*, which appeared in the March, 1943, issue of this journal.

section 1 again, let us add another interval $d\lambda$ after $d\eta$, containing a point λ, and denote the total affective stimulation at λ, reckoning rewards as positive and punishments negatively, by $V(\lambda)$. In human beings this may be determined psychophysically (Horst, 22), and in animals held constant and fitted indirectly from the data. We may designate by $\Delta_{\xi\eta\lambda}\, T_{ij}(\delta, t)$ the change in the ordinary increment $\Delta_{\xi\eta}\, T_{ij}(\delta, t)$ of T_{ij}, as given by (6), which results from the facts: (1) that at $d\xi$, there is type i stimulation of intensity $P_i(\xi)$; (2) at $d\eta$, there is conditioned type j response tendency derived from conditioning of magnitude $Q_j(\eta)$; while (3) at $d\lambda$ there is a quantity of affective stimulation $V(\lambda)$. Then we shall say

$$\Delta_{\xi\eta\lambda}\, T_{ij}(\delta, t) = V(\lambda)\, P_i(\xi)\, Q_j(\eta) \cdot e^{-\alpha\delta-\beta^2(\eta-\xi-\delta)^2-\gamma(t-\lambda)-\varepsilon(\lambda-\eta)}.$$

The considerations leading up to the factor $P_i(\xi)\, e^{-\alpha\delta-\beta^2(\eta-\xi-\delta)^2-\gamma(t-\lambda)}$ in this are the same as in (6), with the exception that here we have $\gamma(t-\lambda)$ instead of $\gamma(t-\eta)$ as before; this results from the fact that the increment (14) is not established until λ, and consequently cannot begin to decay until then. The factor $e^{-\varepsilon(\lambda-\eta)}$ arises from the necessity of some monotone decreasing factor, involving the distance $\lambda-\eta$, which is unity for zero separation and approaches zero asymptotically for increasing $\lambda-\eta$; it may perhaps be rendered plausible on the basis of some hypothesis involving the stimulus-trace of the conditioned part $Q_j(\eta)$, considered as a medium whereby the value of V at λ influences $T_{ij}(\delta, t)$ in a manner involving the value of Q_j at η. That the function of the distance $\lambda-\eta$ must have these properties follows from the experimentally observed diminution of the effect of affective stimulation as the interval between it and the reaction is augmented. The reason for using Q_j here in place of the perhaps more natural E_j, or possibly S_j, are partially observational: firstly, instrumental counter-conditioning by use of punishment can be set up in eductions of the reaction by the conditioned stimulus alone, in which case S_j vanishes throughout (Pavlov, 37 and Youtz, 55, 56);—this excludes the use of S_j for Q_j in (14)—moreover, we observe in general a resistance to instrumental counterconditioning which appears greater when the unconditioned reaction is very strong than it would seem to be under (14) with E_j for Q_j, for in this case the inhibition would be correspondingly more rapid in its rate of increase. This is of course not conclusive, but it appears to confer a presumption. A better founded decision must await more quantitative experimental evidence.

The expression (16) may be added to (5) and integrated over all intervals $d\xi \to d\eta \to d\lambda$ which are concerned; and we obtain, upon a change in order of integration,

$$T_{ij}(\delta, t) = T^0_{ij}(\delta, t) + \int_0^t [S_j(\eta) - Q_j(\eta)]\, d\eta \int_0^\eta P_i(\xi) \times e^{-\alpha\delta-\beta^2(\eta-\xi-\delta)^2-\gamma(t-\eta)}\, d\xi + \int_0^t Q_j(\eta)\, d\eta \int_\eta^t V(\lambda) \times d\lambda \int_0^\eta P_i(\xi)\, e^{-\alpha\delta-\beta^2(\eta-\xi-\delta)^2-\gamma(t-\lambda)-\varepsilon(\lambda-\eta)}\, d\xi\,. \tag{17}$$

Equations (1) and (2) yield with this, if we remember (3), and again apply Dirichlet's rule,

$$Q_j(t) = \int_0^t \left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\| [S_j(\eta) - Q_j(\eta)]\, d\eta + \int_0^t \left\| \begin{matrix} \alpha, \beta, \gamma \\ \varepsilon, \eta, t \end{matrix} \right\| Q_j(\eta)\, d\eta\,, \tag{18}$$

where we have set

$$\left\| \begin{matrix} \alpha, \beta, \gamma \\ \varepsilon, \eta, t \end{matrix} \right\| = \int_\eta^t d\theta \int_\eta^t V(\lambda)\, d\lambda \int_0^\eta F(\xi, \theta) \times e^{-\alpha(t-\theta)-\beta^2(\eta-\xi-t+\theta)^2-\gamma(\theta-\lambda)-\varepsilon(\lambda-\eta)}\, d\xi\,. \tag{19}$$

Equation (18), like (10), is an integral equation to determine $Q_j(t)$, and, by (3) and (17), also E_j and T_{ij}. Since the same continuity conditions are satisfied as before, it is soluble by the standard methods, a process we may carry out in exact analogy to the solution of (10). Then we may put

$$K(t, \eta) = \left\| \begin{matrix} \alpha, \beta, \gamma \\ \varepsilon, \eta, t \end{matrix} \right\| - \left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\|,$$

$$K^{(1)}(t, \eta) = K(t, \eta), \quad K^{(n+1)}(t, \eta) = \int_\eta^t K(\eta, \zeta) K^{(n)}(\zeta, t)\, d\zeta\,,$$

for all integers n, and

$$\Phi(t, \eta) = \sum_{n=1}^\infty K^{(n)}(t, \eta)\,.$$

Then similar manipulation to that of the previous case yields

$$Q_j(t) = \int_0^t \left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\| S_j(\eta)\, d\eta + \int_0^t \int_0^\eta \Phi(t, \eta) \left\| \begin{matrix} \alpha, \beta, \gamma \\ \theta, \eta \end{matrix} \right\| S_j(\theta)\, d\theta\, d\eta\,. \tag{20}$$

The computation of this solution in an actual case would be rather laborious, especially since the indicated quadratures in the expressions (9) and (19) for $\left\| \begin{matrix} \alpha, \beta, \gamma \\ \eta, t \end{matrix} \right\|$ and $\left\| \begin{matrix} \alpha, \beta, \gamma \\ \varepsilon, \eta, t \end{matrix} \right\|$ cannot be performed in closed form. We shall consider approximate solutions in a later study. Finally, we may remark that if the reaction considered is subject also to experimental extinction (55), the amount of this effect may be calculated from (15), but using the new value of Q_j.

5. The case of related stimuli and reaction-tendencies.

In the present section we shall sketch very briefly the extension of the results of previous sections to more general cases, where the values of functions involving one type of stimulus or reaction must be regarded as influencing the course of the learning process in quite different, although related, stimulus-response configurations. There is observational evidence to indicate that not only must association between stimuli—which results in generalization and transfer and the phenomena of discrimination—be taken into account, but also the relations between reactions, and, as a consequence of the latter, between reaction-thresholds. For this latter, e.g., Pavlov (37) has remarked that inhibition— which we treat as a rise in the reaction-threshold—is often generalized very widely, sufficiently so in fact to produce a sleep-like state in the organism. By proceeding in the same way as above, we shall be able to construct a theory capable of accounting for at least the comparative magnitudes of effects of the group mentioned.

We shall consider the various stimulus-and-reaction configurations involved in the situation as point-vectors in a psychophysical n-space, whose mutual distances may be taken for learning purposes as measuring the extent of their association. With human subjects we may locate the stimulus-and-reaction configurations in an appropriate vector space by the standard methods of modern psychophysics and factor analysis, scaling them so as to obtain their mutual distances in terms of "similarity" or "likeness," as estimated by the experimental population.

We shall suppose there to be some N stimulus-vectors $\bar{x}_i [i = 1, 2, \cdots, N]$, and M response-vectors, not necessarily in the same space as the $\bar{x}_i$, of such a kind that, if any stimulus-vector $\bar{x}$ has an unconditioned tendency to evoke a reaction-vector $\bar{y}$, this is the case solely because of its contiguity to some among the point-vectors $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_N$, which have the property of evoking reaction-tendencies for some of the reaction-vectors $\bar{z}_1, \bar{z}_2, \cdots, \bar{z}_M$ which are contiguous to $\bar{y}$. In particular, if we define $T_o(\bar{x}, \bar{y}, \delta, t)$ as the amount of unconditioned response-tendency having the configuration $\bar{y}$ that a unit intensity of the stimulus-vector $\bar{x}$, presented at the time t, will evoke at $t + \delta$, we shall suppose that

$$T_o(\bar{x}, \bar{y}, \delta, t) = \sum_{i=1}^{N} \sum_{j=1}^{M} T_o(\bar{x}_i, \bar{z}_j, \delta, t)\, e^{-k^2[(\bar{x}_j - \bar{x})^2 + (\bar{z}_j - \bar{y})^2]}. \tag{21}$$

If $T(\bar{x}, \bar{y}, \delta, t)$ be the quantity of which $T_o(\bar{x}, \bar{y}, \delta, t)$ is the unconditioned part, we shall consider the above expression to reflect a more general fact: namely, that any increment $\Delta T(\bar{x}, \bar{y}, \delta, t)$ in the total value of the conditioning $T(\bar{x}, \bar{y}, \delta, t)$ for given vectors $\bar{x}$, $\bar{y}$ will bring to the value of $T(\bar{z}, \bar{w}, \delta, t)$ for like configurations $\bar{z}$, $\bar{w}$ an increment of strength

$$\Delta_{\bar{x}, \bar{y}} T(\bar{z}, \bar{w}, \delta, t) = \Delta T(\bar{x}, \bar{y}, \delta, t)\, e^{-k^2[(\bar{x} - \bar{z})^2 + (\bar{y} - \bar{w})^2]}. \tag{22}$$

Here, and correspondingly in (18), $(\bar{x} - \bar{z})^2$, $(\bar{y} - \bar{w})^2$ represent the squares of the magnitudes of the vectors $\bar{x} - \bar{z}$ and $\bar{y} - \bar{w}$, and are equal to the distances between the points whose radii vectores are respectively $\bar{x}$ and $\bar{z}$, or $\bar{y}$ and $\bar{w}$. The cases in the previous sections differ from our present ones in that all the stimuli and reactions occurring in the former are assumed to be infinitely far apart. Our choice of the normal distribution function to express the diminution of generalization with distance is consequent upon its simplicity as compared with other functions of the coordinates of the stimulus point-vectors which are unity with vanishing argument and approach zero asymptotically with increasing distance; but it may perhaps be also rendered plausible on the basis of a neurological hypothesis. There is evidence to indicate that generalization gradients are empirically at least rather like this: thus we may instance the results of Hovland (23), (24), and of Bass and Hull (1).

Let us denote by $\bar{x}(t)$ the stimulus configuration being presented at the time t; and by $P(t)$ the intensity of that stimulation. When $P(t)$ is zero, so that nothing is being presented, it will not matter what value is assigned to $\bar{x}(t)$. Further, let $E(\bar{y}, t)$ denote the intensity of the response-tendency at the time t to react with the response-configuration $\bar{y}$; the unconditioned part of this is

$$S(\bar{y}, t) = \int_0^t P(\theta) \sum_{i=1}^{N} \sum_{j=1}^{M} T_o(\bar{x}_i, \bar{z}_j, t - \theta, \theta)\; e^{-k^2[(\bar{x}_i - \bar{x}(\theta))^2 + [\bar{z}_j - \bar{y}]^2}\, d\theta. \tag{23}$$

As in (2), we shall also have

$$E(\bar{y}, t) = \int_0^t P(\theta)\, T[x(\theta), y, t-\theta, \theta]\, d\theta . \tag{24}$$

We may now repeat an argument substantially like that leading up to (7), but taking into account (22) and a limiting process with respect to associated configurations. We derive

$$T(\bar{y}, \bar{z}, \delta, t) = T_o(\bar{y}, \bar{z}, \delta, t) + \int_V e^{-k^2(\bar{w}-\bar{y})^2}\, dV$$
$$\int_0^t [S(\bar{w}, \eta) - Q(\bar{w}, \eta)]\, d\eta \int_0^\eta P(\xi) \times \tag{25}$$
$$e^{-a\delta-\beta^2(\eta-\xi-\delta)^2-\gamma(t-\eta)-k^2[\bar{x}(\xi)-\bar{y}]^2}\, d\xi ,$$

where the first integration on the right is a volume integral taken throughout the whole psychophysical space in which the response-configurations are located, $\bar{w}$ is the vector from the origin to the element of integration dV, and $Q_j = E_j - S_j$. Again, as before, we have from this

$$Q_j(\bar{y}, t) = \int_V e^{-k^2(\bar{w}-\bar{y})^2}\, dV \int_0^t [S(\bar{w}, \eta) - Q(\bar{w}, \eta)]\, d\eta \times$$
$$\int_\eta^t d\theta \int_0^\eta P(\xi)\, P(\theta) \tag{26}$$
$$e^{-a(t-\theta)-\beta^2(\eta-\xi-t+\theta)^2-\gamma(t-\eta)-k^2[\bar{x}(\xi)-\bar{x}(\theta)]^2}\, d\xi .$$

From this expression, precisely as before, $Q(\bar{y}, t)$ may be determined, although the practical derivation of the solution may be very difficult, since the volume integral in (26), if written out in terms of the coordinates of $\bar{w}$, would involve n successive integrations from $-\infty$ to ∞. The equation would be of mixed Fredholm-Volterra type.

We may proceed similarly to find the expression for $R(\bar{x}, t)$, the reaction-threshold for the response-configuration $\bar{x}$ at the time t, of which $R_o(\bar{x}, t)$ is the unconditioned part. Abbreviating

$$\left\| \begin{matrix} \alpha, \beta, \gamma, t \\ k, w, \eta, y \end{matrix} \right\| = \int_\eta^t d\theta \int_0^\eta P(\xi)\, P(\theta)$$
$$e^{-a(t-\theta)-\beta^2(\eta-\xi-t+\theta)^2-\gamma(y, t-\eta)-k^2\{[\bar{x}(\xi)-\bar{x}(\theta)]^2+(\bar{w}-\bar{y})^2\}}\, d\xi ,$$

we shall have, analogous to (15),

$$R(\bar{y}, t) = R_0(\bar{y}, t) + \int_0^t d\eta \int_V \left\| \begin{matrix} \alpha, \beta', \gamma', t \\ k, \bar{w}, \eta, \bar{y} \end{matrix} \right\| [S(\bar{w}, \eta) - Q(\bar{w}, \eta)]\, dV. \tag{27}$$

Finally, if reward and punishment be introduced, we shall have, giving $V(\lambda)$ the same meaning as before,

$$Q(\bar{y}, t) = \int_0^t d\eta \int_V \left\| \begin{matrix} \alpha, \beta, \gamma, t \\ k, w, \eta, y \end{matrix} \right\| S(\bar{w}, \eta)\, dV + \int_0^t d\eta \int_V Q(\bar{w}, \eta)\, dV \Bigg\{ - \left\| \begin{matrix} \alpha, \beta, \gamma, t \\ k, w, \eta, y \end{matrix} \right\| + \int_\eta^t d\theta \int_\eta^t V(\lambda)\, d\lambda \int_0^\eta P(\xi)\, P(\theta)\, e^{-\alpha(t-\theta)-\beta^2(\eta-\xi-t+\theta)^2-\gamma(\theta-\lambda)-\varepsilon(\lambda-\eta)-k^2\{[\bar{x}(\xi)-\bar{x}(\theta)]^2+(\bar{w}-\bar{y})^2\}}\, d\xi \Bigg\}. \tag{28}$$

On both of these equations, (27) and (28), V is the whole n-space, and $\bar{w}$ is again the radius vector to dV. For evidence, incidentally, that the situation of equation (28) actually occurs, i.e., that instrumental conditioning is in fact generalized, we may mention, among others Youtz (57), Münzinger and Dove (36), and Thorndike (51).

It is perhaps worthwhile to remark that certain previous theories of generalization, notably that of K. W. Spence, (48) and the elaboration thereupon by Hull (28), which have been strikingly confirmed in cases of "relational" versus "absolute" transfer as a basis for various phenomena in conditioned discrimination, emerge as special cases of the theoretical account given above, except that in accord with more recent data on the question, we agree with Razran (41) in preferring a negatively accelerated gradient to Spence's parabola. This difference is a minor one, however, and consequently the evidence of Gullikson (15), Gullikson and Wolfle (16), and Spence himself (47) in support of these hypotheses may also be used in favor of the present account.

6. Application to experimental situations.

The precise relation of the quantities we have calculated in the previous section to observation is something which must be stated in rather greater detail than hitherto. We may, in general, distinguish three distinct types of situation with regard to the interpretation of our results: (1) The first of these may be represented by the conditioning experiment: here the subject may make any one of a number of mutually exclusive types of response, or, if his reaction-tendency for more of them is high enough, he may make no reaction of the group at all. (2) In the second type of situation, the subject is required to choose one among a certain group of mutually exclusive responses, so that failure to manifest any of them is not a possible alternative. For this type, we may choose as reference situation the case of the rat, which, when placed upon the platform of the learning apparatus, may jump either to the right or to the left, cannot do both, but must do one of them. (3) Thirdly, we may consider the case, properly falling under (1), where the response is not an all-or-none affair, but permits of quantitative variation in magnitude. Examples of this may be found in conditioning glandular or vasomotor reactions to various stimuli.

Suppose we define the *effective response-tendency* for responses of a given type j as

$$\mathcal{E}_j(t) \equiv E_j(t) - R_j(t). \tag{29}$$

Now, in a situation of the first type considered above, let the effective reaction-tendencies at a given time t be $\mathcal{E}_j(t)\,[j = 1, \cdots, M]$. Consider the quantities $K_j(t) = \mathcal{E}_j(t) - \sum_{i \neq j} \mathcal{E}_i(t)$. At most, one of these quantities can be positive during a given interval. If none are, then we shall say that no reaction will occur during the interval. If, however, during an interval $t, t + \delta$, a given $K_j > 0$, then we shall say that a type j response will occur after a reaction-latency from t given by

$$L_j = m \log \frac{\log K_j - q}{\log K_j - (q + r)}, \tag{30}$$

for suitable constants m, q, r. The equation (27) is one derived by Landahl for reaction-times from neurological considerations (see Rashevsky, 40); it appears to fit well to observations. This assertion gives us an account of the observed interference of competing reaction-tendencies at the same time which seems to be in accord with the principal facts in the literature about the subject: see, e.g., Hilgard and Marquis (18), Hull (27), and investigations such as that of Kellogg and Wolf (33).

It is worth remarking that the statement above, together with other similar ones in the present section, does not constitute a new assumption over and above those we have made before; it has, in fact, the character of a definition. It fixes, although not completely, the meaning of the functions E_j and R_j which we introduced above without defining them formally, trusting to the intuition of the reader to supply them with enough meaning to enable him to follow the arguments adduced from experiment. For a theory of quasi-definitions of this sort (they are called *reduction-sentences*, in that they enable us to reduce certain assertions involving the definienda—although not all, as would be the case with explicit definitions—to sentences which are directly confirmable by observation), we refer the reader to various works of R. Carnap, in particular (8) and (9). Quasi-definitions of this kind are very common in science: see, e.g., the physical definition of "electric intensity."

In the second type of situation mentioned, we clearly cannot use the same method of predicting responses: if all the K_j are negative, we infer from the above that nothing will be evoked; but by hypothesis, even in this case some reaction must occur. In this case, we shall regard the subliminal reaction-tendencies $\mathcal{E}_j$ as representing probabilities of the corresponding types of response; and we shall say that, if at any time t the subject is required to make a response of one of several mutually exclusive types $1, 2, \cdots, M$, then the probability of a response of a given type j is

$$P_j = \frac{\mathcal{E}_j}{\sum_{i=1}^{M} \mathcal{E}_i}. \tag{31}$$

The meaning of the $\mathcal{E}_j$ in this case seems to be rather different from that in type (1) situations; but there does not appear to be a simple way of remedying this, since any attempt to represent the $\mathcal{E}_j$ as probabilities in type (1) situations encounters the difficulty of providing a plausible value for the probability of no response. Moreover, type (2) situations can hardly avoid being much more under voluntary control; and by this route a number of disregarded factors may become important, to hold which constant a statistical procedure of gathering data and interpreting our formulae may be made necessary.

Situations of the third type considered above present a somewhat different sort of problem. We have so far tended to regard $\mathcal{E}_j$ as possessing the dimensions of an average strength of neural stimulation, and the relation of such strength of excitation to the intensity of resulting responses capable of varying magnitudes is not very well known. Extremely tentatively, however, we may perhaps assume the relationship to be a linear one: $I_j = k\,\mathcal{E}_j$ for fairly small positive values of $\mathcal{E}_j$, or else $I_j = k(1 - e^{-\mu \mathcal{E}_j})$ for positive $\mathcal{E}_j$ and suitable k, μ. The whole matter requires further experimental study.

By way of conclusion we may make several remarks. First, our introductory purpose must be considered fulfilled, since we have devised a theory of sufficient generality to provide an experimental solution for all those cases of learning and conditioning which we set out to consider. Second, our theory agrees qualitatively with a sufficient number of experimentally observed effects to make a *prima facie* case for its being, if not true, at least not far from the truth at many points, so that it should merit careful experimental study. In fact, of the well-known observational effects, there seem to be only two which we can say at present are clearly in disagreement with the theory: namely, that the facility of experimental extinction is apparently much reduced when courses of extinction have been interpolated in the conditioning process; and that the process of "second-order conditioning," in which a third stimulus is conditioned to a reaction by use, in place of an unconditioned stimulus, of one previously conditioned to the reaction, although rather difficult to bring about, nevertheless occurs, whereas under our theory this process would only produce conditioned inhibition—unless affective elements are involved, as in the experiments of Brogden, Lippman, and Culler (3), and Brogden and Culler (2), where we should derive the proper prediction. We shall attempt to account for such cases, and in addition examine the possibilities for grounding the theory neurologically, in a later study, where we shall also make quantitative comparison of the theory with whatever precise data are available, and attempt to approximate some of the more complex and involved functions in the theory in a simpler manner. Meanwhile, the reader is invited to compare the theory with any suitable experimental results.