JOURNAL OF MATHEMATICAL PSYCHOLOGY 14, 130–143 (1976)

Optional Shift, and Discrimination Learning with Redundant Relevant Dimensions: Predictions of an Attentional Model

M. FRANK NORMAN

Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19174

The Zeaman–House–Lovejoy (ZHL) model for discrimination learning involves interacting perceptual and response learning processes, with rate parameters φ and θ, respectively. This paper describes a number of predictions of the ZHL model for the optional shift paradigm. Kendler and Kendler showed that the proportion of subjects who learn an optional shift as a reversal increases with subject's age, and Campione and others have argued that this trend may reflect developmental variation in parameters like φ and θ. Our main result is a formula specifying the dependence of reversal probability on φ and θ for the ZHL model.

1. INTRODUCTION

The most influential attentional model for discrimination learning is the one formulated by Zeaman and House (1963). A similar scheme was proposed by Lovejoy (1966). A previous paper (Norman, 1974) surveyed the predictions of the Zeaman–House–Lovejoy (or ZHL) model for a variety of experimental paradigms. The present paper continues this survey. The topics to be covered are optional shift and learning with redundant relevant dimensions. The notation, style, and viewpoint of the present and previous papers are similar, and together they treat a substantial proportion of the standard discrimination learning paradigms.

TABLE 1

An Optional Shift Experiment[a]

Initial phase	Shift phase	Test phase
(Gc^+, Rt^-)	(Gc^-, Rt^+)	(Gt^+, Rc^+)[b]
(Gt^-, Rc^+)		

[a] G = green, R = red, t = triangle, c = circle, $+$ = rewarded, $-$ = nonrewarded.
[b] Choice of Gt is indicative of reversal shift.

An optional shift experiment has three phases, illustrated in Table 1. Stimuli vary along two dimensions, form, and color. Each dimension is represented by two values: circle and triangle, and green and red, respectively. On any trial, the subject (a young child, perhaps) must choose between a green circle and a red triangle [such a trial is symbolized (*Gc*, *Rt*)] or between a green triangle and a red circle [(*Gt*, *Rc*)]. The spatial position of the two stimuli within each pair is randomized over trials. Rewarded and nonrewarded choices are indicated by "+" and "−" superscripts.

In the Initial Phase, both types of stimulus presentations occur equally often, and choices of circular stimuli are rewarded. We say that form is *relevant*, color is *irrelevant*, and circle is *correct*. After a criterion is met in the Initial Phase, the subject enters the Shift Phase. Here only one of the stimulus pairs, (Gc^-, Rt^+), is presented, and choice of the red triangle is rewarded. We say that color and form are *redundant relevant dimensions*. The subject can solve the problem by learning to choose "red" or "triangle" (or "red triangle"). In the first case, the solution in the Shift Phase is based on the dimension which was irrelevant in the Initial Phase, so the subject has performed an *extradimensional shift*. In the second case, the solution is based on the initially relevant dimension, but the subject has had to learn a new correct response. This is *reversal shift*.

The Test Phase permits us to diagnose the basis of solution of the shift problem. When confronted with (*Gt*, *Rc*), a subject who performed a reversal and learned to choose triangle in the Shift Phase will presumably choose *Gt* consistently. On the other hand, consistent choice of *Rc* indicates that the subject performed an extradimensional shift and learned to choose red in the Shift Phase. Subjects who choose neither stimulus consistently in the test phase are termed *nonselectors*. Both stimuli of the *test pair*, (*Gt*, *Rc*), are rewarded in the Test Phase. Interspersed with presentations of the test pair are presentations of the *relearning pair*, (Gc^-, Rt^+), which occurred in the Shift Phase.

Kendler and Kendler (1962) have been interested for many years in developmental changes in discrimination learning performance. Recently (Kendler & Kendler, 1970) they performed a large scale optional shift experiment on children (grades *K*, 2, and 6) and college students that provides very clear documentation for such changes. Their main finding is this: With increasing age, the proportion of reversal shifters increases, while the proportions of extradimensional shifters and nonselectors decrease. In the Kendlers' view, this developmental trend is indicative of the appearance of verbal mediation, or some other form of mediation that facilitates reversal shift, in larger proportions of subjects in older subject groups. In accordance with this outlook, they propose that qualitatively different models be applied to children before and after they attain the crucial mediational capacities. This program seems quite reasonable, intuitively.

It has, however, been recently noted that the developmental trend reported by the Kendlers is susceptible to another interpretation. According to this view, all subjects undergo perceptual and response learning in accordance with some two-process attentional scheme, but the relative rates of perceptual and response learning change with age (Campione, 1970; Dickerson, Novik, & Gould, 1972; Zeaman & House, 1974). The verbal demonstration (see Campione, 1970, p. 300) that learning rate parameters control reversal probability in attentional theories seems perfectly sound, but it gives no indication of the magnitude of behavioral change that will result from a given degree of parametric variation. Alternatively, the verbal argument provides no basis for estimating the parametric variation corresponding to the observed developmental trend. We shall fill this gap, at least for the ZHL model, by deriving a formula relating the probability of optional reversal shift to the perceptual learning rate parameter, φ, and the response learning rate parameter, θ. This formula is Eq. (1) of the next section.

2. The ZHL Model

The following description of the ZHL model applies to all three phases of the experiment. There is a perceptual learning process, indexed by the probability, v, that the subject attends to form on a particular trial. If he does not attend to form, he attends to color; this has probability $v' = 1 - v$. There are two response learning processes, one for each dimension. If the subject attends to form on a trial, he chooses circle with probability y, and triangle with probability y'. If he attends to color, he chooses green with probability z and red with probability z'. In summary,

$$\begin{aligned} P(\text{form}) &= v, \\ P(\text{color}) &= v' = 1 - v, \\ P(\text{circle} \mid \text{form}) &= y, \\ P(\text{green} \mid \text{color}) &= z. \end{aligned}$$

The three variables, v, y, and z, or, alternatively, the vector $x = (v, y, z)$, determine the subject's state of learning at the beginning of any trial.

The values of v, y, and z change from trial to trial as a function of the perceptual response (form or color), the overt response, and the consequent reward or non-reward. The probability, v, of attending to form increases if the subject attends to form and is rewarded, or attends to color and is not rewarded. Otherwise v decreases. The conditional probability, y, of choosing circle, given attention to form, changes only on trials on which the subject attends to form. If he attends to form, chooses circle, and is rewarded, y increases. The probability y also increases when the subject attends to form, chooses triangle, and is not rewarded. On other trials on which the subject attends to form, y decreases. Changes in z are analogous to the changes in y just described.

All of these changes are effected by linear operators. Let V_n, Y_n, and Z_n be the values of v, y, and z on trial n of some phase of the experiment, $n = 0, 1, 2,...$, and let $\Delta V_n = V_{n+1} - V_n$. Then increases and decreases in V_n are described by equations of the form

$$\Delta V_n = \varphi V_n' \quad \text{and} \quad \Delta V_n = -\varphi V_n,$$

where $0 < \varphi \leqslant 1$. Similarly, the possible changes in Y_n and Z_n are

$$\Delta Y_n = \theta Y_n' \quad \text{and} \quad \Delta Y_n = -\theta Y_n,$$

$$\Delta Z_n = \theta Z_n' \quad \text{and} \quad \Delta Z_n = -\theta Z_n,$$

where $0 < \theta \leqslant 1$. Thus φ is the perceptual learning rate parameter, while θ is the response learning rate parameter. We shall see that the ratio, φ/θ, of the two learning rates is a crucial determinant of performance in optional shift experiments.

The Initial Phase of the experiment is the basic discrimination learning paradigm, and was considered in detail in Norman (1974). For our present purposes, we need only note that, according to the model, each subject learns to attend to the relevant, form, dimension (hence $V_n \to 1$ as $n \to \infty$, almost surely [a.s.]) and the correct, circle, response on this dimension is completely learned ($Y_n \to 1$ as $n \to \infty$, a.s.). Thus, if the criterion in the Initial Phase is sufficiently strict, we may take $v = 1$ and $y = 1$ to be the initial values of v and y in the Shift Phase for all subjects. Since reward is random with respect to color in the Initial Phase, the natural choice for the initial value of z in the Shift Phase is $z = \frac{1}{2}$. However, this choice is only an approximation to the appropriate value, which may vary from subject to subject, depending on initial color preferences and the particular sequence of responses and rewards on trials on which the subject attends to the irrelevant color dimension.

We now turn our attention to the Shift Phase, which will be the focus of our subsequent work. Table 2 describes the trial-to-trial changes in v, y, and z that can

TABLE 2

The ZHL Model for Experiments with Redundant Relevant Dimensions[a]

Event	Δv	Δy	Δz	Probability
form, t^+	$\varphi v'$	$-\theta y$	0	vy'
form, c^-	$-\varphi v$	$-\theta y$	0	vy
color, R^+	$-\varphi v$	0	$-\theta z$	$v'z'$
color, G^-	$\varphi v'$	0	$-\theta z$	$v'z$

[a] t = triangle, c = circle, R = Red, G = Green, $+$ = rewarded, $-$ = nonrewarded.

occur in this phase, and gives their probabilities. This description does not depend on the initial values of v, y, and z, so Table 2 applies to any experiment with two redundant relevant dimensions, not just to the Shift Phase of an optional shift experiment. As an illustration of the content of the table, consider the first row, which corresponds to trials on which the subject attends to form and chooses triangle. According to the table, this event has probability $V_n Y_n'$, and produces changes

$$\Delta V_n = \varphi V_n', \qquad \Delta Y_n = -\theta Y_n, \qquad \Delta Z_n = 0,$$

in v, y, and z on trial n of the Shift Phase.

We pause here to give an overview of our main results for this model.

THEOREM 1. *V_n converges, as $n \to \infty$, to a limit, V_∞, and $V_\infty \in \{0, 1\}$* a.s. *Subjects who learn to attend to form ($V_\infty = 1$) also learn to choose triangle ($Y_n \to 0$); similarly, $V_\infty = 0$ implies $Z_n \to 0$* a.s.

This theorem holds for arbitrary values of V_0, Y_0, and Z_0, though the probability, $P(V_\infty = 1)$, that a subject learns to attend to form is influenced by the initial conditions, as well as the learning rates φ and θ. For the particular initial conditions $V_0 = 1$, $Y_0 = 1$, $Z_0 = \frac{1}{2}$, $P(V_\infty = 1) = P(\text{reversal shift})$ is the probability that a subject learns an optional shift as a reversal. We shall derive the formula

$$\begin{aligned} P(\text{reversal shift}) &\doteq 1 - 2^{-1}\varphi/\theta \qquad \text{if} \quad \varphi/\theta \leqslant 2, \\ &\doteq 0 \qquad\qquad\qquad \text{if} \quad \varphi/\theta \geqslant 2. \end{aligned} \tag{1}$$

The symbol "$\doteq$" denotes approximate equality. The approximation should be good when φ and θ are small, since it becomes exact in the limit as φ and $\theta \to 0$. Even when φ and θ are not small, (1) may be a useful rough guide to the relation between P(reversal shift) and φ/θ.

According to (1), P(reversal shift) can take on any value between 0 and 1. It is less than $\frac{1}{2}$ if $\varphi/\theta > 1$, and greater than $\frac{1}{2}$ if $\varphi/\theta < 1$. Moreover P(reversal shift) would increase with age, as observed by Kendler and Kendler (1970), if φ/θ decreased with age. Independent evidence of such a change in the ratio of perceptual and response learning rates has been provided by Dickerson *et al.* (1972) and Medin (1973, p. 333). [See Zeaman and House (1974) for a thorough assessment of the current status of theories bearing on developmental trends in discrimination learning.]

The fact that the ZHL model is compatible with any observed value of P(reversal shift) raises the specter of a model that is consistent with any data, hence correct but useless. [For a criticism of this sort in a closely related context, see T. Kendler (1971, last paragraph on p. 770 and first paragraph on p. 771).] Let us immediately put this apparition to rest. The fact that $V_\infty \in \{0, 1\}$ a.s. means that, according to the model, each subject will be classified as either a reversal shifter or an extra-dimensional shifter in the Test Phase; i.e., there will be no nonselectors, provided that the criterion in the Shift Phase is sufficiently strict to permit the learning process in this Phase to approach asymptote. This prediction is clearly at odds with Kendler and Kendler's (1970) data, so the ZHL model cannot be considered a complete account of learning in their experiment. Readers are invited to devise modifications of the model or of the experiment that will bring the two into closer agreement.

3. Convergence of V_n

The purpose of this section is to prove Theorem 1 and Theorem 2 below. These theorems pertain to experiments with redundant relevant dimensions, hence to the Shift Phase of the optional shift experiment.

For $n = 0, 1, 2, \ldots$, let

$$v_n = E(V_n), \qquad y_n = E(Y_n), \qquad z_n = E(Z_n).$$

In applications to the Shift Phase, $n = 0$ corresponds to the initial trial of that phase. Let T be the total number of errors (choices of Gc^-) in the experiment or phase under consideration.

THEOREM 2.

$$P(V_\infty = 1) = v_\infty = v_0 + (\varphi/\theta)[(z_0 - z_\infty) - (y_0 - y_\infty)], \tag{2}$$

$$E(T) = \theta^{-1}[(y_0 - y_\infty) + (z_0 - z_\infty)]. \tag{3}$$

In applications to the Shift Phase, we set $v_0 = 1$, $y_0 = 1$, and $z_0 = \frac{1}{2}$. However, y_∞ and z_∞ are unknown, so Theorem 2 does not, in itself, show how $P(V_\infty = 1)$ and $E(T)$ depend on φ and θ. Approximations to y_∞, z_∞, $P(V_\infty = 1)$, and $E(T)$, valid when φ and θ are small, are developed in the next section.

Referring back to Table 2, we note the strong and psychologically unwarranted assumption that there is only a single "φ" parameter, instead of four, and a single θ parameter, instead of four. Theorem 1 remains true if this assumption is dropped, though the proof of the more general theorem is more difficult than that given below. Theorem 2 and the theorems presented in the next section are easily generalized to the case where there are two θ's, one for y and one for z, but only one φ.

The following simple lemma is the basis for all our subsequent work. Let $x = (v, y, z)$ and $X_n = (V_n, Y_n, Z_n)$.

LEMMA 1.

$$\varphi^{-1}E(\Delta V_n \mid X_n = x) = -vy + v'z, \tag{4}$$

$$\theta^{-1}E(\Delta Y_n \mid X_n = x) = -vy, \tag{5}$$

$$\theta^{-1}E(\Delta Z_n \mid X_n = x) = -v'z. \tag{6}$$

Proof. From Table 2 we obtain

$$\varphi^{-1}E(\Delta V_n \mid X_n = x) = v'vy' - v^2y - vv'z' + (v')^2z.$$

Equation (4) is obtained by rewriting v^2y as $vy - vv'y$ and $(v')^2z$ as $v'z - vv'z$, and noting that the four terms with vv' as a factor sum to zero. The derivations of (5) and (6) are even more straightforward. Q.E.D.

Proofs of Theorems 1 *and* 2. Let

$$W_n = \varphi^{-1}V_n - \theta^{-1}Y_n + \theta^{-1}Z_n. \tag{7}$$

It follows from (4), (5), and (6) that

$$E(\Delta W_n \mid X_n) = 0. \tag{8}$$

Hence $\{W_n, n \geqslant 0\}$ is a martingale. Since it is bounded ($|W_n| \leqslant \varphi^{-1} + \theta^{-1}$), $W_\infty = \lim_{n\to\infty} W_n$ exists a.s. (Chung, 1974, Theorem 9.4.4). But $Y_{n+1} \leqslant Y_n$ and $Z_{n+1} \leqslant Z_n$ so $Y_\infty = \lim_{n\to\infty} Y_n$ and $Z_\infty = \lim_{n\to\infty} Z_n$ exist. Convergence of W_n, Y_n, and Z_n implies convergence of V_n, i.e., $V_\infty = \lim_{n\to\infty} V_n$ exists.

According to Table 2, $\Delta V_n = \varphi V_n'$ or $\Delta V_n = -\varphi V_n$. In the first case,

$$|\Delta V_n| = \varphi V_n' \geqslant \varphi V_n' V_n.$$

In the second case

$$|\Delta V_n| = \varphi V_n \geqslant \varphi V_n V_n'.$$

Hence, in either case, $|\Delta V_n| \geqslant \varphi V_n V_n'$. Letting $n \to \infty$ in this inequality, and noting that $\Delta V_n \to V_\infty - V_\infty = 0$, we obtain $V_\infty(1 - V_\infty) = 0$. Hence $V_\infty \in \{0, 1\}$ a.s.

By (5),

$$\begin{aligned} \theta^{-1}E(\Delta Y_n) &= \theta^{-1}E(E(\Delta Y_n \mid X_n)) \\ &= -E(V_n Y_n). \end{aligned} \tag{9}$$

Adding over $n = 0, 1, \ldots, m-1$, we obtain

$$\theta^{-1}(y_0 - y_m) = \sum_{n=0}^{m-1} E(V_n Y_n).$$

This yields

$$\theta^{-1}(y_0 - y_\infty) = \sum_{n=0}^{\infty} E(V_n Y_n), \tag{10}$$

on letting $m \to \infty$. In particular,

$$\infty > \sum_{n=0}^{\infty} E(V_n Y_n) = E\left(\sum_{n=0}^{\infty} V_n Y_n\right),$$

so $\sum_{n=0}^{\infty} V_n Y_n < \infty$ and $V_\infty Y_\infty = 0$ a.s. Now

$$[V_\infty = 1] - [Y_\infty = 0] \subset [V_\infty Y_\infty > 0],$$

so the event on the left has probability 0. If $P(A - B) = 0$, we say that A implies B a.s. Hence $V_\infty = 1$ implies $Y_\infty = 0$ a.s. Similarly, $V_\infty = 0$ implies $Z_\infty = 0$ a.s. This completes the proof of Theorem 1.

Since $V_\infty \in \{0, 1\}$ a.s., it follows that $v_\infty = E(V_\infty) = P(V_\infty = 1)$, as asserted in (2). To derive the second part of (2), take expectations on both sides of (8) to obtain $E(\Delta W_n) = 0$, hence $\Delta E(W_n) = 0$. It follows that $E(W_\infty) = E(W_0)$, which, in conjunction with the definition, (7), of W_n yields (2).

Adding (10) to the analogous equation,

$$\theta^{-1}(z_0 - z_\infty) = \sum_{n=0}^{\infty} E(V_n' Z_n),$$

for Z_n, we obtain

$$\theta^{-1}[(y_0 - y_\infty) + (z_0 - z_\infty)] = \sum_{n=0}^{\infty} E(V_n Y_n + V_n' Z_n). \tag{11}$$

But $V_n Y_n + V_n' Z_n$ is the conditional probability of an error on trial n, given X_n, so its expectation is the unconditional probability of an error on trial n. It follows that the sum on the right in (11) is $E(T)$, and (3) is proved. Q.E.D.

4. Slow Learning

In this section we consider the model of Table 2 with initial values V_0, Y_0, and Z_0 subject only to the condition that their variances are zero. Thus there are constants v_0, y_0, and z_0 such that

$$V_0 = v_0, \qquad Y_0 = y_0, \qquad Z_0 = z_0, \tag{12}$$

a.s. We shall develop approximations to v_∞, y_∞, z_∞, and $E(T)$ that become exact in the limit as φ and θ approach zero. The initial values v_0, y_0, and z_0 are held fixed as φ and θ vary. For example, the natural initial values for the Shift Phase, $v_0 = 1$, $y_0 = 1$, and $z_0 = \frac{1}{2}$, do not depend on φ and θ.

To derive the desired approximations, we begin by taking expectations over $X_n = x$ on both sides of (5) to obtain

$$\theta^{-1}\Delta E(Y_n) = -E(V_n Y_n).$$

If we replace $E(V_n Y_n)$ by the approximation $E(V_n)\, E(Y_n) = v_n y_n$, we obtain

$$\theta^{-1}\Delta y_n \doteq -v_n y_n . \tag{13}$$

If the approximate equality sign is changed to exact equality, the resulting equation, together with similar equations obtained from (4) and (6), defines the ***expected operator approximation*** to v_n, y_n, and z_n. We shall go a step further, and replace the difference equation, (13), by a differential equation. Let $v(t)$, $y(t)$, and $z(t)$ be functions of a continuous variable, t, whose values, $v(n\theta)$, $y(n\theta)$, and $z(n\theta)$, at time $n\theta$ approximate v_n, y_n, and z_n. Then (13) implies that

$$\theta^{-1}(y(t+\theta) - y(t)) \doteq -v(t)\, y(t),$$

where $t = n\theta$. The left-hand side is approximately $(dy/dt)(t)$. This suggests that we *define* $y(t)$ to be the solution of the differential equation

$$dy/dt = -v(t)\, y(t). \tag{14}$$

To be more precise, we must first write the analogous equations for dv/dt and dz/dt derived from (4) and (6):

$$dv/dt = (\varphi/\theta)[-v(t)\, y(t) + v(t)'\, z(t)], \tag{15}$$

$$dz/dt = -v(t)'\, z(t), \tag{16}$$

where $v(t)' = 1 - v(t)$. The functions $v(t)$, $y(t)$, and $z(t)$ are then defined to be the solutions of the system (14)–(16) of differential equations satisfying the initial conditions

$$v(0) = v_0, \qquad y(0) = y_0, \qquad z(0) = z_0, \tag{17}$$

analogous to (12). Note that φ and θ do not appear in (14) or (16), and only their ratio appears in (15). This implies that $v(t)$, $y(t)$, and $z(t)$ depend on φ and θ only through their ratio φ/θ.

A very general theorem (Norman, 1972, Theorem 8.1.1.B) implies that $v(t)$, $y(t)$, and $z(t)$ are uniquely defined by (14)–(17), and that

$$v_n \to v(t), \qquad y_n \to y(t), \qquad z_n \to z(t), \tag{18}$$

as $n \to \infty$, $\varphi \to 0$, and $\theta \to 0$, in such a way that $n\theta \to t < \infty$ and φ/θ is constant. It is but a short step, conceptually, to Theorem 3, which is the result that we need for our present purposes. In (18), $n \to \infty$ and $\theta \to 0$ simultaneously. In Theorem 3, $\theta \to 0$ after $n \to \infty$.

THEOREM 3. *Let* $v_\infty = \lim_{n\to\infty} v_n$, *let* $v(\infty) = \lim_{t\to\infty} v(t)$, *and let* y_∞, z_∞, $y(\infty)$, *and* $z(\infty)$ *be defined similarly. Then*

$$v_\infty \to v(\infty), \qquad y_\infty \to y(\infty), \qquad z_\infty \to z(\infty),$$

as $\varphi \to 0$ *and* $\theta \to 0$ *in such a way that* φ/θ *remains fixed.*

This result permits us to use $v(\infty)$, $y(\infty)$, and $z(\infty)$ to approximate v_∞, y_∞, and z_∞ when φ and θ are small. The utility of this approximation lies in the fact that $v(\infty)$, $y(\infty)$, and $z(\infty)$, unlike v_∞, y_∞, and z_∞, are easy to calculate. It is, unfortunately, necessary to distinguish between the cases $y_0 > z_0$, $y_0 = z_0$ and $y_0 < z_0$.

THEOREM 4. *Suppose that* $y_0 > z_0$. (a) *If* $v_0 > (\varphi/\theta)(y_0 - z_0)$, *then* $z(\infty) = 0$, $y(\infty) = 0$, *and*

$$v(\infty) = v_0 - (\varphi/\theta)(y_0 - z_0) > 0.$$

(b) *If* $v_0 \leqslant (\varphi/\theta)(y_0 - z_0)$, *then* $z(\infty) = 0$, $v(\infty) = 0$, *and*

$$y(\infty) = (y_0 - z_0) - (\theta/\varphi)v_0.$$

THEOREM 5. *Suppose that* $y_0 = z_0$. *Then* $v(\infty) = v_0$, $y(\infty) = 0$, *and* $z(\infty) = 0$.

The case $y_0 < z_0$ can be treated by applying Theorem 4 to $\tilde{V}_n = 1 - V_n$, $\tilde{Y}_n = Z_n$, and $\tilde{Z}_n = Y_n$.

The following corollary specializes Theorem 4 to the initial conditions of the Shift Phase.

COROLLARY 1. *Suppose that* $v_0 = 1$, $y_0 = 1$, *and* $z_0 = \frac{1}{2}$.

(a) *If* $\varphi/\theta < 2$, *then* $z(\infty) = 0$, $y(\infty) = 0$, *and*

$$v(\infty) = 1 - 2^{-1}\varphi/\theta.$$

(b) *If* $\varphi/\theta \geqslant 2$, *then* $z(\infty) = 0$, $v(\infty) = 0$, *and*

$$y(\infty) = 2^{-1} - (\theta/\varphi).$$

Combining Corollary 1 with Theorem 3, we obtain Eq. (1) of Section 2. The next result combines Corollary 1 and Theorem 2. Recall that T is the total number of errors.

COROLLARY 2. *Suppose* $v_0 = 1$, $y_0 = 1$, *and* $z_0 = \frac{1}{2}$. *Then*

$$\begin{aligned} E(T) &\sim \tfrac{3}{2}\theta^{-1} && \text{if} \quad \varphi/\theta \leqslant 2, \\ &\sim \theta^{-1} + \varphi^{-1} && \text{if} \quad \varphi/\theta \geqslant 2, \end{aligned} \tag{19}$$

as $\varphi \to 0$, $\theta \to 0$, *and* φ/θ *remains fixed.*

Were one in possession of a set of optional shift data with a negligible proportion of nonselectors, one could use (1) and (19) to estimate both φ and θ.

Theorem 5 applies to an experiment with redundant relevant dimensions and equal initial response biases y_0 and z_0 ($y_0 = z_0 = \frac{1}{2}$ is the most intuitive special case). According to Theorem 5, $y(\infty) = z(\infty) = 0$; hence, by Theorem 3, $y_\infty \doteq 0$ and $z_\infty \doteq 0$ when φ and θ are small. In other words, response learning goes almost to completion on both dimensions, not just the one on which the subject happens to be absorbed. The possibility of "multiple-cue learning" by the ZHL model was emphasized by Shepp, Kemler, and Anderson (1972).

It remains only to prove Theorems 3–5. Theorems 4 and 5 must be proved before Theorem 3.

Proof of Theorem 4. It follows from (18) that $0 \leqslant v(t), y(t), z(t) \leqslant 1$. By (14), $dy/dt \leqslant 0$, so $y(\infty)$ exists. Similarly, (16) implies that $z(\infty)$ exists.

Comparing (15) with (14) and (16), we see that

$$dv/dt = (\varphi/\theta)[(dy/dt) - (dz/dt)].$$

Integration yields

$$v(t) - v_0 = (\varphi/\theta)[(y(t) - y_0) - (z(t) - z_0)].$$

Hence $v(\infty)$ exists and

$$v(\infty) - v_0 = (\varphi/\theta)[(y(\infty) - y_0) - (z(\infty) - z_0)]. \tag{20}$$

Letting $t \to \infty$ in (14), we see that

$$\lim_{t\to\infty} dy/dt = -v(\infty)\, y(\infty).$$

The only value of this limit that is compatible with the boundedness of y is 0, hence

$$v(\infty)\, y(\infty) = 0. \tag{21}$$

Similarly, (16) yields

$$v(\infty)'\, z(\infty) = 0. \tag{22}$$

Suppose now that $y_0 > z_0$.

Case 1. Suppose that $v(\infty) > 0$. Then $y(\infty) = 0$, by (21), and (20) yields

$$v(\infty) - v_0 \leqslant (\varphi/\theta)(z_0 - y_0) < 0.$$

But $v_0 \leqslant 1$, so $v(\infty) < 1$, and (22) yields $z(\infty) = 0$. Hence, by (20),

$$v(\infty) = v_0 - (\varphi/\theta)(y_0 - z_0) > 0,$$

and $v_0 > (\varphi/\theta)(y_0 - z_0)$. The latter inequality is the defining condition of Case (a) of Theorem 4.

Case 2. Suppose $v(\infty) = 0$. By (22), $z(\infty) = 0$, so (20) yields

$$\begin{aligned} -v_0 &= (\varphi/\theta)[(y(\infty) - y_0) + z_0] \\ &\geqslant (\varphi/\theta)(z_0 - y_0), \end{aligned}$$

or $v_0 \leqslant (\varphi/\theta)(y_0 - z_0)$, as in Case (b) of Theorem 4.

Consideration of both cases shows that $v(\infty) > 0$ if and only if $v_0 > (\varphi/\theta)(y_0 - z_0)$. In other words, Cases (a) and (b) of Theorem 4 coincide, respectively, with Cases 1 and 2 of the proof. Reviewing the proof with this in mind, we see that all assertions of the theorem have been established. Q.E.D.

Proof of Theorem 5. Equations (21) and (22) are still valid when $y_0 = z_0$, and (20) reduces to

$$v(\infty) - v_0 = (\varphi/\theta)(y(\infty) - z(\infty)). \tag{23}$$

We consider three cases.

Case 1. Suppose that $0 < v(\infty) < 1$. Then $y(\infty) = z(\infty) = 0$ by (21) and (22), so $v(\infty) = v_0$ by (23).

Case 2. Suppose that $v(\infty) = 1$. Then $y(\infty) = 0$ by (21), so (23) yields

$$1 - v_0 = -(\varphi/\theta)\, z(\infty).$$

Since $1 - v_0 \geqslant 0$ and $-(\varphi/\theta)\, z(\infty) \leqslant 0$, we must have $v_0 = 1 = v(\infty)$, and $z(\infty) = 0$. A similar argument shows that $v(\infty) = v_0$ and $y(\infty) = z(\infty) = 0$ if $v(\infty) = 0$ (Case 3). Thus these conclusions hold in any case. Q.E.D.

Proof of Theorem 3. Suppose first that $y_0 \geqslant z_0$. By Theorems 4 and 5, $z(\infty) = 0$, consequently $z_\infty \geqslant z(\infty)$. Similarly, $y_\infty \geqslant 0 = y(\infty)$ if $y_0 = z_0$ or if $y_0 > z_0$ and $v_0 > (\varphi/\theta)(y_0 - z_0)$. If $y_0 > z_0$ and $v_0 \leqslant (\varphi/\theta)(y_0 - z_0)$, consider (2) in the previous section. This equation implies that

$$0 \leqslant v_0 + (\varphi/\theta)[(z_0 - y_0) + y_\infty],$$

or

$$\begin{aligned} y_\infty &\geqslant (y_0 - z_0) - (\theta/\varphi)v_0 \\ &= y(\infty), \end{aligned}$$

by (b) of Theorem 4. Thus we have shown that $y_\infty \geqslant y(\infty)$ and $z_\infty \geqslant z(\infty)$ whenever $y_0 \geqslant z_0$. By symmetry,

$$y_\infty \geqslant y(\infty) \qquad \text{and} \qquad z_\infty \geqslant z(\infty), \tag{24}$$

for any values of y_0 and z_0.

Let t be any positive real number and let $n = [t/\theta]$, the largest integer that does not exceed t/θ. Then $n \to \infty$ and $n\theta \to t$ as $\theta \to 0$, so (18) is applicable. But y_n (like Y_n) is a nonincreasing sequence, so $y_\infty \leqslant y_n$, and (18) implies that $\limsup_{\theta\to 0} y_\infty \leqslant y(t)$. Since this holds for all $t > 0$, we must have $\limsup_{\theta\to 0} y_\infty \leqslant y(\infty)$. This result, in combination with (24), yields $\lim_{\theta\to 0} y_\infty = y(\infty)$. The proof that $\lim_{\theta\to 0} z_\infty = z(\infty)$ is similar. (We assume, as usual, that φ/θ is fixed as θ and φ approach 0.) Letting θ approach 0 in (2), we see that

$$\begin{aligned} \lim_{\theta\to 0} v_\infty &= v_0 + (\varphi/\theta)[(y(\infty) - y_0) - (z(\infty) - z_0)] \\ &= v(\infty), \end{aligned}$$

by (20). Q.E.D.

Acknowledgments

I am indebted to D. W. Evans and D. G. Kemler for their help on this project. I am also grateful to a reviewer for calling my attention to Dickerson, Novik, and Gould (1972) and Zeaman and House (1974).

References

Campione, J. C. Optional intradimensional and extradimensional shifts in children as a function of age. *Journal of Experimental Psychology*, 1970, **84**, 296–300.
Chung, K. L. *A course in probability theory*. 2nd ed. New York: Academic Press, 1974.

DICKERSON, D. J., NOVIK, N., & GOULD, S. A. Acquisition and extinction rates as determinants of age changes in discrimination shift behavior. *Journal of Experimental Psychology*, 1972, **95**, 116–122.

KENDLER, H. H., & KENDLER, T. S. Vertical and horizontal processes in problem solving. *Psychological Review*, 1962, **69**, 1–16.

KENDLER, T. S., & KENDLER, H. H. An ontogeny of optional shift behavior. *Child Development*, 1970, **41**, 1–27.

KENDLER, T. S. Review of *Mechanisms of animal discrimination learning*, by N. S. Sutherland and N. J. Mackintosh. *Contemporary Psychology*, 1971, **16**, 769–771.

LOVEJOY, E. Analysis of the overlearning reversal effect. *Psychological Review*, 1966, **73**, 87–103.

MEDIN, D. L. Subproblem analysis of discrimination shift learning. *Behavior Research Methods and Instrumentation*, 1973, **5**, 332–336.

NORMAN, M. F. *Markov processes and learning models*. New York: Academic Press, 1972.

NORMAN, M. F. Effects of overtraining, problem shifts, and probabilistic reinforcement in discrimination learning: predictions of an attentional model. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology*, Vol. 1. San Francisco: Freeman, 1974. Pp. 185–208.

SHEPP, B. E., KEMLER, D. G., & ANDERSON, D. R. Selective attention and the breadth of learning: an extension of the one-look model. *Psychological Review*, 1972, **79**, 317–328.

ZEAMAN, D., & HOUSE, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 159–223.

ZEAMAN, D., & HOUSE, B. J. Interpretations of developmental trends in discriminative transfer effects. In Anne D. Pick (Ed.), *Minnesota symposia on child psychology*, Vol. 8. Minneapolis: Univ. of Minnesota Press, 1974. Pp. 144–186.

RECEIVED: July 29, 1975