Reprinted from
The British Journal of Mathematical and Statistical Psychology,
Vol. 21, Part 2, p. 193, November 1968

DISCRIMINATIVE BINARY PREDICTION WITH REINFORCED CUE IDENTIFICATION

By Dominic W. Massaro
University of California, San Diego

and John W. Moore
University of Massachusetts, Amherst

Two tones, T_1 and T_2, differing in loudness, were employed as stimuli in a discriminative binary prediction task. Two responses, A_1 and A_2, predicting events E_1 and E_2 respectively, were available to the subjects. An A_1 response was correct with probability π_1 and π_2 on T_1 and T_2 trials, respectively. The present study assessed the effects of the covariation of π_2 and the similarity of T_1 and T_2 on $P(A_1|T_1)$, the probability of an A_1 response given a T_1 trial. Subjects were also required to identify the trial type (loud or soft) and were given partial feedback of identification responding. Partial feedback was sufficient to eliminate the effects of cue similarity on discriminative event prediction and also improved identification of trial types. The reinforcement effects indicated that subjects learn to behave appropriately to identified cues rather than learn to make specific motor responses.

1. Introduction

The experimental situation of discriminative binary prediction consists of a series of trials, each of which commences with the onset of one of two cue stimuli, T_1 or T_2. The probability of T_1 is β, and the probability of T_2 is $1-\beta$. Two responses, A_1 and A_2, predicting events E_1 and E_2 respectively, are available to the subject. On a T_1 trial, an A_1 response is correct with probability π_1, and an A_2 response is correct with probability $1-\pi_1$. On a T_2 trial, an A_1 response is correct with probability π_2, and an A_2 response is correct with probability $1-\pi_2$.

A previous experiment by Massaro *et al.* (1968) has shown that $P(A_1|T_1)$, the probability of an A_1 response given a T_1 trial, is dependent upon both π_2 and the similarity between T_1 and T_2. With values of $\pi_1=0{\cdot}8$ and $\pi_2=0{\cdot}8$, $0{\cdot}5$ or $0{\cdot}2$, the results indicated that $P(A_1|T_1)$ was: (1) a linear function of π_2 when the cues were highly similar, (2) independent of π_2 and equal to π_1 at an intermediate level of cue similarity, and a U-shaped function of π_2 when the cues were highly discriminable. In that study, in order to determine the similarity between T_1 and T_2, the subjects were required to indicate which of the two cues (loud or soft tones) was being presented before making their prediction of the event that would occur on that trial. However, the subjects were not given feedback regarding the accuracy of their identification of trial type. The present experiment is an assessment of the effects of variations of π_2 and cue similarity on $P(A_1|T_1)$ with partial feedback (25 or 75 per cent) about identification responses.

2. Method

Subjects The subjects were 144 University of Massachusetts undergraduates and they were assigned randomly to the experimental treatments.

Apparatus Up to four subjects were run at a time, each seated at a tabletop enclosure containing an Estes Straughan conditioning board, which consisted of a white centre warning light and two red event lights, each positioned above a spring-loaded lever switch. The centre of the board contained a $2 \times 4{\cdot}5$ in. panel recessed 1·5 in. with two spring-loaded buttons mounted one above the other and labelled loud and soft. Tones were generated by a Hewlett–Packard Model 200 audio-oscillator and were presented over matched headphones with a continuous white masking noise. Experimental events were controlled by Lehigh Valley 1620 Probability Randomizers, Hunter Interval Timers, and relays. Events and responses were recorded on an Esterline–Angus event recorder.

Procedure The onset of a tone started a trial. The tone lasted 2·5 sec., during which the subjects were required to make a loud or soft identification response by pressing the respective button. At the end of the tone, the correct button was illuminated for 0·9 sec. on feedback trials. On both feedback and no feedback trials, the warning light was illuminated 0·9 sec. after tone offset. The light was left on for 1·5 sec. and the subjects made their prediction responses. The event light was illuminated for 0·9 sec. immediately following the end of the warning light. Hence each trial lasted 5·8 sec. The inter-trial interval was 4·2 sec. Each subject received 300 trials in which all stimuli, feedback and events were presented randomly by an appropriate setting on the probability randomizers.

Subjects were given the following instructions: 'You will be receiving two tones differing slightly in loudness over the headphones. Your first task will be to indicate which of the tones you are listening to. You will do this by pressing one of the two buttons. You will press the top button for the louder tone and the bottom button for the softer tone. Notice the buttons are labelled loud and soft. You are expected to guess if you are not sure of your decision. You will have 2½ sec. to make your choice. After 2½ sec., the tone will go off and one of the two buttons *may* come on, indicating whether the loud or soft tone did come on. Immediately after this, the white light on top of your panel will go on. Your second task is to *predict* which of the two red lights at the bottom of your panel will come on. As soon as possible after the white light comes on, you are to flip one of the two switches. After you have flipped a switch, one of the two red lights will show. If the red light above the switch you pressed comes on, you were correct. If the other one comes on, you were incorrect on that trial. Remember, you will always be told whether or not you have been correct on the second task. You will not always be given information regarding your indication of a loud or soft tone.'

Design Three levels of cue similarity (ΔI) were crossed with three stimulus contingent reinforcement schedules and two feedback conditions. The two feedback conditions were 25 and 75 per cent feedback. The intensity pairings of the two 800 Hz tones were 73–74·5, 73–76 and 73–79 db SPL, giving a 1·5, 3 and 6 db intensity differential (ΔI), respectively, for three decreasing levels of cue similarity. The three reinforcement schedules (π) were: (*a*) $\pi_1 = P(E_1|T_1) = 0{\cdot}8$ for all groups, and (*b*) $\pi_2 = P(E_1|T_2) = 0{\cdot}8$, 0·5 and 0·2 respectively for group 0·8–0·8, group 0·8–0·5 and group 0·8–0·2. Thus, there were eight subjects in each cell of a $3 \times 3 \times 2$ between-groups factorial design for a total of 144 subjects. The analysis of variance of $P(A_1)$'s included the three between variables of ΔI, π and feedback, and the three within variables of cue (T_1 vs. T_2), feedback trial (yes or no), and trial block (three blocks of 100 trials).

3. Results

Marginal statistics Fig. 1 shows that discrimination performance, $P(A_1|T_1) - P(A_1|T_2)$ ($F = 180{\cdot}21$; d.f. $= 1, 126$; $P < 0{\cdot}001$), increased as the difference between π_1 and π_2 increased ($F = 81{\cdot}48$; d.f. $= 2, 126$; $P < 0{\cdot}001$). Discrimination performance also increased over training ($F = 19{\cdot}56$; d.f. $= 2, 252$) and increased faster over training for larger π_1 and π_2 differences ($F = 7{\cdot}00$; d.f. $= 4, 252$; $P < 0{\cdot}001$). The cue $\times\, \Delta I$ interaction and their 15 higher order interactions were not statistically significant. This finding that cue similarity did *not* disrupt discrimination performance in any significant way indicates that partial feedback was sufficient to eliminate the effects of cue similarity found in previous studies (Massaro *et al.*, 1968).

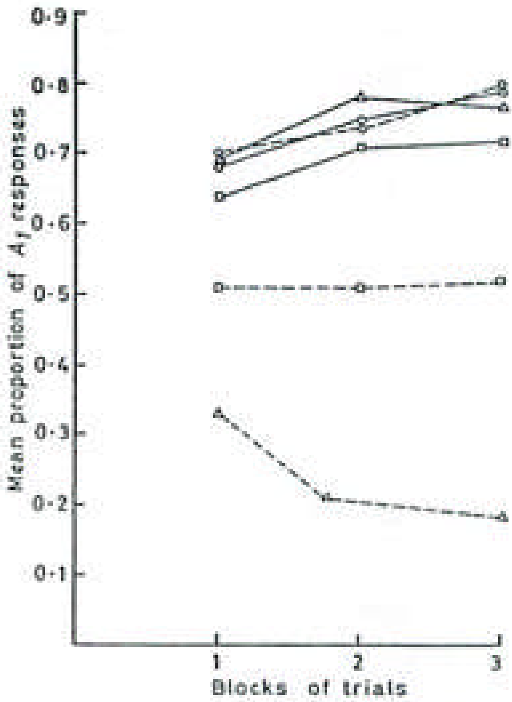

FIGURE 1. Mean proportion of A_1 responses on T_1 and T_2 trials as a function of trial blocks and π_2. ——, $P(A_1|T_1)$; – – – –, $P(A_1|T_2)$; ○, 0·8–0·8 group; □, 0·8–0·5; △, 0·8–0·2.

Fig. 1 also shows that, overall, $P(A_1|T_1)$ was a U-function of π_2 ($F = 172{\cdot}7$; d.f. $= 1, 126$; $P < 0{\cdot}001$). The fact that $P(A_1|T_1)$ was about 5 per cent lower when $\pi_2 = 0{\cdot}5$ than when $\pi_2 = 0{\cdot}8$ or $0{\cdot}2$ seems to indicate that a discriminative stimulus with little or no cue value (i.e. one reinforced on a 0·5–0·5 basis) depressed $P(A_1|T_1)$ relative to a stimulus with cue value. This finding agrees with studies where the trial types were highly discriminable (Popper & Atkinson, 1958), except that the overshooting reported in that study was not found in the present experiment. This overall depression of $P(A_1|T_1)$ could be due to the randomness of the occurrence of feedback in the present study. For example, the subjects may have hypothesized that event occurrence was also dependent on whether feedback was given on a trial. This hypothesized contingency would give subjects many more strategies to test, which would decrease differential responding and eliminate any tendency to overshooting.

The analyses of identification responses indicated that correctness of identification increased with decreases in cue similarity ($F = 44{\cdot}05$; d.f. $= 2, 126$; $P < 0{\cdot}001$), and that the number of correct identifications increased over training ($F = 15{\cdot}73$; d.f. $= 2, 252$; $P < 0{\cdot}001$). The mean percentage of correct identifications for $\Delta I = 1{\cdot}5$, 3 and 6 db were 72, 86 and 95 per cent, respectively. The mean percentage of correct identifications of an earlier study (Massaro *et al.*, 1968) for $\Delta I = 1{\cdot}5$, 3 and 6 db were 63, 79 and 93 per cent, respectively. The fact that no feedback was presented in the earlier study indicates that the partial feedback presented in this study improved identification responding significantly, especially with the more similar cues.

Sequential statistics Massaro (1969) has developed a three-state Markov model of discriminative probability learning for reinforcement schedules where $\pi_1 > 0{\cdot}5$ and $\pi_2 < 0{\cdot}5$. Fig. 2 illustrates the possible transitions among the conditioning states: appropriate (A), unconditioned (U) and inappropriate (I). For example, suppose the subject is in state U. A T_1 trial will occur with probability β.

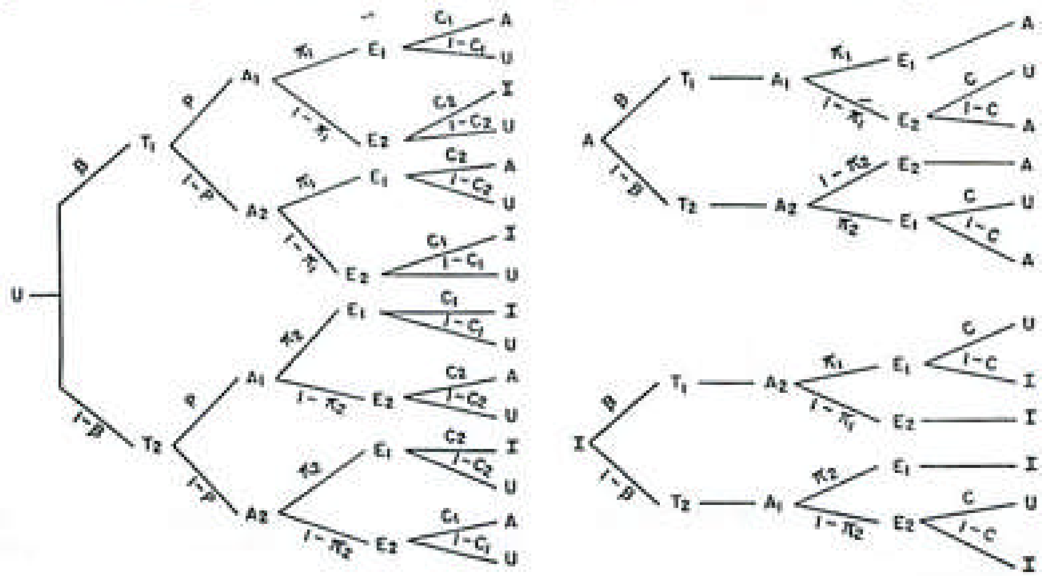

FIGURE 2. Branching process, starting from each state on trial n, for the model when $\pi_1 > 0{\cdot}5$ and $\pi_2 < 0{\cdot}5$.

By the response axiom, the subject will make an A_1 response with probability p which will be reinforced by an appropriate event with probability π_1. Therefore, with probability c_1, the subject goes into state A, and with probability $1 - c_1$, he remains in state U. These transitions among conditioning states lead to the following transition matrix P and response probability vectors for appropriate responses, $P(A_1|T_1)$ and $P(A_2|T_2)$:

$$P = \begin{array}{c} \\ A \\ U \\ I \end{array}\!\!\begin{array}{c} \begin{array}{ccc} A & U & I \end{array} \\ \begin{bmatrix} 1 - c(1-\pi_a) & c(1-\pi_a) & 0 \\ E & 1 - E - F & F \\ 0 & \pi_a c & 1 - \pi_a c \end{bmatrix} \end{array} \qquad \begin{array}{c} P(A_1|T_1) \\ \begin{bmatrix} 1 \\ p \\ 0 \end{bmatrix} \end{array} \qquad \begin{array}{c} P(A_2|T_2) \\ \begin{bmatrix} 1 \\ 1-p \\ 0 \end{bmatrix} \end{array} \tag{1}$$

where

$$\pi_a = \beta\pi_1 + (1-\beta)(1-\pi_2),$$
$$E = \beta\pi_1\{pc_1 + (1-p)c_2\} + (1-\beta)(1-\pi_2)\{pc_2 + (1-p)c_1\},$$
$$F = \beta(1-\pi_1)\{pc_2 + (1-p)c_1\} + (1-\beta)\pi_2\{pc_1 + (1-p)c_2\}.$$

The basic assumption of the theory is that subjects are reinforced for behaving appropriately to a distinctive cue rather than for making a specific motor response. That is, subjects are reinforced for making an appropriate response in the sense of one having the highest likelihood of being correct on a particular stimulus trial. For example, if the right-hand light is the most frequent event, given the trial type, and the subject is reinforced for making a right-hand prediction, he is also reinforced for making an appropriate response. Therefore he will be more likely to make the same response on the next trial, provided it is of the same type. But if the next trial is different, the subject will still be more likely to make the appropriate response, which means a response that is physically opposite from the response on the previous trial. Thus, the subject is reinforced not for a particular (right or left) event prediction but for an appropriate (most frequently correct) or inappropriate response. The theory has accounted for results of experiments that have included variations of π, β and cue similarity (Massaro, 1969).

Another test of the model in the present study is the prediction of reinforcement effects when responses are analysed in terms of prior identification response rather than trial type. Table 1 shows the observed and predicted values of the first-order conditional probabilities for group 0·8–0·2 (25 per cent feedback condition) on no feedback trials when subjects correctly identified the trial type. In the estimation of the parameters, c_1 was found to be equal to zero. This result indicates that a subject could leave state U only if he was incorrect (i.e. $c_1 = 0$). This fact is appealing, since the states represent different strategies and it seems likely that subjects will change strategies only when they are incorrect. In addition, p was assumed to be equal to $P(E_1) = \beta\pi_1 + (1-\beta)\pi_2$. The two parameters, c and c_2, were estimated using a minimum χ^2 criterion between predicted and observed values for the 16 independent first-order conditional probabilities presented in the Appendix. The data were pooled over all trials to increase the total number of observations. The predicted values were obtained by letting the probability vector at the start of trial 1 be

$$w_1 = [0 \quad 1 \quad 0],$$

and repeatedly computing the vector w_n for every trial n by the equation

$$w_n = w_{n-1}P,$$

where P is the transition matrix defined by eqn. (1). Then the entries of the average probability vector were taken as the values for the corresponding probabilities of the three states in the expressions for the conditionals presented in the Appendix.

TABLE 1. OBSERVED AND PREDICTED VALUES OF $P(A_{1,n+1}|T_{i,n+1}T_{j,n}A_{k,n}E_{l,n})$ FOR GROUP 0·8–0·2

(' No feedback ' trials; trial type correctly identified for the 25 per cent feedback condition pooled over levels of ΔI and trial blocks.)

$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{l,n}$	Frequency Observed	Frequency Predicted	Proportion Observed	Proportion Predicted
1	1	1	1	365	380·9	0·867	0·905
1	1	1	2	84	89·0	0·737	0·746
1	1	2	1	59	52·1	0·634	0·560
1	1	2	2	17	15·4	0·436	0·395
1	2	1	1	10	9·9	0·400	0·395
1	2	1	2	63	61·0	0·578	0·560
1	2	2	1	93	91·7	0·756	0·746
1	2	2	2	430	433·4	0·898	0·905
2	1	1	1	42	45·1	0·089	0·095
2	1	1	2	20	34·8	0·146	0·254
2	1	2	1	50	57·2	0·385	0·440
2	1	2	2	15	20·0	0·455	0·605
2	2	1	1	14	13·3	0·636	0·605
2	2	1	2	44	47·1	0·411	0·440
2	2	2	1	30	28·5	0·268	0·254
2	2	2	2	50	45·6	0·104	0·095

$c=0·314$; $c_2=0·334$; $\chi^2=11·20$.

TABLE 2. OBSERVED AND PREDICTED VALUES OF $P(A_{1,n+1}|T_{i,n+1}T_{j,n}A_{k,n}E_{l,n})$ FOR GROUP 0·8–0·2

(Feedback trials; trial type correctly identified for the 75 per cent feedback condition pooled over levels of ΔI and trial blocks.)

$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{l,n}$	Frequency Observed	Frequency Predicted	Proportion Observed	Proportion Predicted
1	1	1	1	623	599·7	0·941	0·906
1	1	1	2	107	104·6	0·728	0·711
1	1	2	1	74	63·9	0·692	0·597
1	1	2	2	12	8·7	0·546	0·394
1	2	1	1	19	10·6	0·704	0·394
1	2	1	2	4	59·7	0·040	0·597
1	2	2	1	93	81·1	0·816	0·711
1	2	2	2	415	412·2	0·912	0·906
2	1	1	1	57	54·8	0·098	0·094
2	1	1	2	28	39·0	0·207	0·289
2	1	2	1	25	38·3	0·263	0·403
2	1	2	2	17	20·0	0·515	0·606
2	2	1	1	21	21·8	0·583	0·606
2	2	1	2	41	47·5	0·348	0·403
2	2	2	1	46	35·2	0·377	0·289
2	2	2	2	48	50·6	0·089	0·094

$c=0·383$; $c_2=0·413$; $\chi^2=76·85$.

The observed rank-orderings in Table 1 agree with the predictions of the model based on the subject making an *appropriate* response. Therefore, when trial types n and $n+1$ are identified as the same,

$$P(A_1|A_1E_1) > P(A_1|A_1E_2) > P(A_1|A_2E_1) > P(A_1|A_2E_2),$$

whereas when the trial types are identified as different, the rank-ordering is exactly opposite. Table 2 shows that reinforcement of identification responses does not disrupt this mode of response, since the rank-orderings are the same as for the trials without feedback. These results support the thesis that, in discrimination learning, subjects are reinforced for behaving appropriately to identified cues rather than for making specific motor responses.

4. Concluding Remarks

In discrimination learning, the reinforcement schedule in operation for one trial type does affect responding to the other trial type. A stimulus with some cue validity enhances appropriate responding to the other stimulus compared to a stimulus with no cue value (i.e. one reinforced on a 0·5–0·5 basis). When the subject is reinforced for behaving appropriately (i.e. predicting the most frequent event given the trial type) he will be more likely to behave appropriately on the following trial.

Acknowledgements

This research was carried out during the senior author's tenure as a National Aeronautics and Space Administration predoctoral fellow under grant NSG(T)–137.

Reprints may be obtained from D. W. Massaro, Department of Psychology, University of California, San Diego, La Jolla, Calif. 92037.

References

Massaro, D. W. (1969). A three state Markov model for discrimination learning. *J. math. Psychol.* (in press).

Massaro, D. W., Halpern, J. & Moore, J. W. (1968). Generalization effects in human discrimination learning with overt cue identification. *J. exp. Psychol.* 77, 474–482.

Popper, J. & Atkinson, R. C. (1958). Discrimination learning in a verbal conditioning situation. *J. exp. Psychol.* 56, 21–25.

APPENDIX

Presented below are the expressions for the conditional statistics of the form $P(A_{1,n+1}|T_{i,n+1}T_{j,n}A_{k,n}E_{l,n})$:

$$P(A_1|T_1T_1A_1E_1) = \frac{A+pU\{c_1+(1-c_1)p\}}{A+pU},$$

$$P(A_1|T_1T_1A_1E_2) = \frac{A\{cp+(1-c)\}+U(1-c_2)p^2}{A+pU},$$

$$P(A_1|T_1T_1A_2E_1) = \frac{Icp+(1-p)U\{c_2+(1-c_2)p\}}{I+(1-p)U},$$

$$P(A_1|T_1T_1A_2E_2) = \frac{(1-p)U(1-c_1)p}{I+(1-p)U},$$

$$P(A_1|T_1T_2A_1E_1) = \frac{U(1-c_1)p^2}{I+pU},$$

$$P(A_1|T_1T_2A_1E_2) = \frac{Icp+pU\{c_2+(1-c_2)p\}}{I+pU},$$

$$P(A_1|T_1T_2A_2E_1) = \frac{A\{cp+(1-c)\}+(1-p)U(1-c_2)p}{A+(1-p)U},$$

$$P(A_1|T_1T_2A_2E_2) = \frac{A+(1-p)U\{c_1+(1-c_1)p\}}{A+(1-p)U},$$

$$P(A_1|T_2T_1A_1E_1) = \frac{U(1-c_1)p^2}{A+pU},$$

$$P(A_1|T_2T_1A_1E_2) = \frac{Acp+pU\{c_2+(1-c_2)p\}}{A+pU},$$

$$P(A_1|T_2T_1A_2E_1) = \frac{I\{cp+(1-c)\}+(1-p)U(1-c_2)p}{I+(1-p)U},$$

$$P(A_1|T_2T_1A_2E_2) = \frac{I+(1-p)U\{c_1+(1-c_1)p\}}{I+(1-p)U},$$

$$P(A_1|T_2T_2A_1E_1) = \frac{I+pU\{c_1+(1-c_1)p\}}{I+pU},$$

$$P(A_1|T_2T_2A_1E_2) = \frac{I\{cp+(1-c)\}+U(1-c_2)p^2}{I+pU},$$

$$P(A_1|T_2T_2A_2E_1) = \frac{Acp+(1-p)U\{c_2+(1-c_2)p\}}{A+(1-p)U},$$

$$P(A_1|T_2T_2A_2E_2) = \frac{(1-p)U(1-c_1)p}{A+(1-p)U}.$$