Probability learning of response patterns[1]

GORDON H. BOWER
STANFORD UNIVERSITY

A 4-light, 4-key probability learning task was altered by having S predict which 2 of the 4 lights would occur each trial. The question is whether S's habit hierarchy is best represented as composed of 4 single-key habits from which two responses are successively selected, or composed of 6 unitary response-pair habits from which one pair is selected per trial. The data favor the latter representation, since the asymptotic proportions of response pairs matched the corresponding light-pair probabilities.

This experiment tries to clarify a conceptual puzzle that arises from a slight alteration of the standard probability learning situation. In the standard situation, on each trial S predicts and then is shown one of n events. As viewed by stochastic models (Estes, 1964; Luce, 1959), S's repertoire in this situation is represented as consisting of n response alternatives (where A_i = prediction of event E_i) and their associated strengths or probabilities. Further it is assumed that the n response probabilities are changed each trial in a manner depending upon the "reinforcing" event. Under suitable circumstances, the model and the data (Estes, 1964) agree that response probabilities will asymptotically match the corresponding event probabilities.

Let us now alter this standard situation by collapsing two trials into one. In our experiment, there were 4 event lights and 4 response keys. But here, two of the 4 lights occur on each trial and S's task is to predict the two lights that will occur. There are 6 pairs of lights and E presents pair ij with probability π_{ij} independent of S's responses.

The puzzle is how best to identify the response alternatives to which any stochastic model is to be applied in this situation. Two opposing identifications of S's "habit profile" are apparent; we shall call these the "singlet" and "doublet" identifications, respectively. According to the singlet approach, we identify only the four single-key responses. The reinforcing events are viewed as changing the strengths (probabilities) of these four responses, and it is supposed that S predicts a pair of events by two successive selections from this 4-habit hierarchy (e.g., according to Luce's (1959) ranking rule). By this view, a single response comes to occur at a frequency dependent upon the marginal probability with which its corresponding light has occurred in the paired reinforcing events.

According to the alternative "doublet" identification, we completely segregate the 6 response pairs, treating them as 6 distinct, "unitary" responses with no transfer of strength between overlapping patterns. Thus, reinforcing event E_{ij} increases only the probability of response pair A_{ij} and not other pairs involving A_i or A_j. According to this view, we may expect the asymptotic probability of response doublet A_{ij} to match π_{ij}, the corresponding doublet-event probability. Contrariwise, by the singlet view, such doublet-event probability matching would be expected only for special reinforcement schedules (e.g., uniform π_{ij}'s). Our experiment was designed to decide between these two identifications.

Method

The Ss were 30 students from an Introductory Psychology course; they were tested individually. They comprised two groups of 15 Ss each; the groups differed in the six π_{ij} values, which are listed in rows 1 and 4 of Table 1. Each S had a different sequence of 384 trials, with the π_{ij} proportions realized exactly in each block of 48 trials. The apparatus was a modified Humphreys board with a white warning light, a row of 4 red event lights, and below a corresponding row of 4 push buttons. E controlled the light pair presented each trial by a 6-position rotary switch which was rotated silently between trials. Another rotary switch allowed E to counterbalance across Ss which keys and lights served as A_1E_1, A_2E_2, etc.

Because the instructions are probably quite important, they follow verbatim:

> "This is an experiment concerned with how people develop expectations about events in their environment. Before you is a board with a white light at the top, four red lights in the middle, and four push-buttons, one underneath each of the red lights. The experiment will be divided into a series of separate trials. Each trial commences when the white light at the top comes on. Three seconds later two of the four red bulbs will come on for two seconds, and the trial will end. Your job is to predict as best you can which two bulbs will light at the end of the trial. You are to indicate your predictions by briefly pushing the button under each of the two red bulbs which you expect will light up at the end of the trial. After the white light comes on, you have exactly three seconds in which to make your two predictions, so you will have to make your decisions quickly. You can see whether your predictions are correct by comparing them with the two bulbs that light at the end of the trial. You should strive to be correct as frequently as possible. In the first few trials you will just be guessing, but as you learn more about what to expect the accuracy of your predictions will improve somewhat. Remember that you are to make two predictions (push two different buttons) each trial after the white light and before the red lights. Do not push the two buttons simultaneously; use the index finger of your preferred hand to push the button of your first choice, release it, then push the button of your second choice."

Results

Although the order of occurrences of the two responses each trial was recorded, analysis revealed no systematic trends in these data. Typically Ss used a "left-to-right" order in pushing the two buttons, independent of the reinforcement probabilities. Accordingly, the order will be ignored in discussing the response pair data.

The results most relevant to our question are the asymptotic proportions of the 6 response pairs. These appear rounded to two decimals in rows 2 and 5 of Table 1; they are group averages taken from the final 48 trials, by which time response proportions appeared stable. The observed proportions match the event proportions (the π_{ij}'s) with remarkable accuracy. Indeed, the overall match is closer than that often obtained in the standard single-response, single-event situation. Inspection of individual Ss' proportions revealed no glaring discrepancies from the group averages in Table 1.

To gauge the decisiveness of the results, one particular singlet model was developed. Let Q_i represent the operator corresponding to single event E_i for the standard model, and let Q_{ij} be the operator corresponding to the double event E_{ij}. Define Q_{ij} as $.5(Q_iQ_j + Q_jQ_i)$. Letting $p_{k,n}$ be the probability that A_k is the first response on trial n, the equation for the Q_{ij} operator is

$$Q_{ij}\, p_{k,n} = (1-\theta)^2\, p_{k,n} + d_{ijk}\, \theta\, (1 - .5\,\theta),$$

where d_{ijk} equals 1 if $k = i$ or j, and equals 0 otherwise. Such operators produce asymptotes of $p_k = .5\,\pi_k$, where π_k is the marginal probability that light k occurs. Using Luce's ranking rule (pick most preferred, eliminate it, choose next preferred) for predicting response pairs, the doublet predictions in rows 3 and 6 of Table 1 are then obtained. These predictions are obviously poor. Generally, the singlet predictions err by being more uniform (equiprobable) than are the observed proportions. An informative discrepancy is the A_{23} doublet for Group A. Event E_{23} never occurs and Ss learn this, whereas the singlet model expects some A_{23} responding because it views the single responses A_2 and A_3 as separately strong.

Asymptotic response-pair proportions conditional upon the preceding response-pair and reinforcement were inspected but will not be reported. Too few observations are spread over the $6^3 = 216$ proportions to provide any reliable conclusions.

Table 1. Event proportions, observed response-pair proportions, and singlet predictions for the two groups of Ss. Response proportions are group averages obtained from the final 48 trials. All proportions are rounded to two decimals.

		Response Pairs					
		12	13	14	23	24	34
Group A	Event prob., π_{ij}	.29	.29	.08	0	.08	.25
	Obsd. prob., p_{ij}	.29	.29	.05	0	.10	.26
	Singlet prediction	.17	.26	.19	.13	.10	.15
Group B	Event prob., π_{ij}	.33	.08	.08	.08	.33	.08
	Obsd. prob., p_{ij}	.33	.07	.08	.09	.36	.07
	Singlet prediction	.27	.08	.17	.13	.27	.08

Discussion

The double-response identifications were clearly appropriate in this experiment; that is, the response pairs came to function as unitary patterns. It is not possible to determine from these data whether a shift from component to patterned responding occurred over the course of learning; possibly a more sensitive experiment could decide this issue.

The patterned responding found here raises questions about experimental limits, whether an experimental continuum can be contrived for moving S from component to patterned responding. The temporal placement of the two "reinforcing" lights with respect to the two predictive responses would appear to be critical. Our situation, which produced patterned responding, gave both lights together after the two responses, and the two responses were not distinguished. But suppose the two targets were separately labelled (say A and B), and the trial sequence were "predict A, then predict B; show target A, then target B". Or a slight variant: "predict A, show target A; predict B, show target B". Each of these variants would probably move S somewhat in the direction of "singlet" responding since they approximate more closely a simple convolution of the standard situation. It seems likely too that S's instructions and the reinforcement conditions (e.g., whether the locations of the A and B targets are dependent or independent) would be important factors biassing S towards singlet or doublet responding. Readers familiar with statistical learning theory will notice the functional parallel between this issue (singlet vs. doublet response identifications) and one characterizing different models of discrimination learning (component vs. pattern stimulus identifications). The two problems are indeed analogous.

The experimental alteration tried here suggests a general format of the Humphreys game: on each trial S is to choose K out of N lights, trying to hit as many as he can of T target lights that will appear (where $N > K > T$). Alternatively S might rank order the N alternatives each trial. Such methods would obtain more information about S's habit hierarchy each trial than does the standard single-response method. Unfortunately, little evidence is presently available to even suggest the general outlines of the stochastic models required to deal with the data from such experiments. Previous research on probability learning has dealt exclusively with the $K = T = 1$ case.

References

Estes, W. K. Probability learning. In A. W. Melton (Ed.), *Categories of human learning.* New York: Academic Press, 1964.

Luce, R. D. *Individual choice behavior.* New York: Wiley, 1959.

Note

1. Supported by a grant (HD-00954) to the author from the National Institute of Child Health and Human Development, and by a grant (NGR-05-020-036) to R. C. Atkinson from the National Aeronautics and Space Administration.