

The utility of event variability in two choice probability learning

JOSEPH HALPERN, MADISON DENGLER AND Z. JOSEPH ULEHLA
UNIVERSITY OF DENVER

Two types of instructions and three levels of event variability were factorally combined in a two-choice probability learning task. The rank order of the marginal response probabilities provided some support for a utility theory. It was suggested that such a utility theory might prove useful for several diverse probability learning situations where overshooting is typically obtained.

A relatively common finding of research concerned with choice behavior is that S's choices of two alternative events roughly match the proportions with which these events occur. This probability matching, however, is by no means a ubiquitous finding as several situations have been described where such behavior is rarely obtained. Specifically, probability over-shooting is found when the simple two-choice probability learning task is modified by (a) differential rewards, (b) instruction that emphasizes the random nature of the reinforcing events or instructions that provide S with the reinforcement schedules and (c) increasing the number of alternative choice responses.

An attempt to describe probability learning behavior that can account for this probability overshooting is attributed to Siegel (1964). Siegel assumes that Ss attempt to maximize a dual process utility function including the utility for a correct choice and the utility for choice variability. Experimental conditions that provide response variability without the need to depart from optimal choices reduce, by some proportion, that amount of utility that must be supplied by the S's choice of least likely events. The theory thus provides an ordered hypothesis of increases in asymptotic response probability as a function of choice variability. The overshooting commonly found in situation "a" is clearly predicted from this utility theory, and if one assumes that sequential independence of choices provides response variability (cf., Erickson, 1966), then the theory can account for the overshooting obtained with situations "b" and "c."

Empirical tests of utility theory provide conflicting evidence. Young (1967), and Messick & Rapoport (1965 a, b) have provided some support for an expected utility notion while Messick (1965) presents evidence in direct opposition to a utility theory. Each of the above studies has involved rather extreme modifications of the simple two-choice task. These modifications have taken the form of increased complexity. The present experiment represents an attempt to provide variability and yet retain the simple two-choice nature of the probability learning task. To this end, the two-choice task involved multiple successive reinforcing events.

Method

Subjects. The Ss were 96 University of Denver undergraduate volunteers randomly distributed into six groups of 16 Ss each.

Apparatus. The rim of a vertically mounted wheel was divided into 40 sections of equal size on which event outcomes were presented. Each outcome was either a triangle or a cross which was either blue or orange, with or without a white dot in its center, placed on a white or black background. Event probabilities were such that the alternatives, triangle, orange, no white dot, and white background, occurred 70% of the time. The events of different dimensions were statistically independent. The wheel was covered and only one section was visible when the wheel was not in motion.

Duplicates of the events appeared next to buttons on S's response panel. Each row of the panel corresponded to an event dimension and each button in the row corresponded to one of the two within-dimension alternatives. Centered in each row was a small indicator light.

Procedure. Each trial began with the spinning of the wheel and the onset of one of the response panel indicator lights. The particular light which was illuminated indicated to S the relevant event dimension, and S was instructed to predict which of the alternatives within this dimension would occur. An electric motor spun the wheel for an average of 6 sec during which time Ss made their choice response. Exactly 7.5 sec after trial onset, the reinforcing event on the wheel was illuminated for a 5 sec period.

Up to four Ss were run at a time. The room was darkened to the extent that Ss could not see the reinforcing event prior to its illumination, nor could they observe the activities of adjacent Ss. Groups 1, 2, and 3 were required to make a total of 300 successive two-alternative choices within 1, 2, and 4 event dimensions respectively. In no case was S required to predict alternatives for more than one dimension on a single trial. Event dimensions were equally distributed over Ss in Group 1 and randomly assigned to experimental sessions for Group 2.

Half the Ss in all groups were given reinforcement instructions (condition RI). They were told the exact proportion of each of the relevant alternative events (28:12), and the wheel was made visible by removal of the cover for the session. Ss in the NI condition were given no information about reinforcement proportions.

To eliminate positional response bias, the buttons on each row of the response panel represented each alternative equally often. Pilot studies established that Ss

express no consistent preference for either of the alternatives within any of the dimensions.

Results and Discussion

Curves showing the proportion of A_1 responses in blocks of 30 trials for each condition and group are presented in Fig. 1. The figure shows clear learning curves for Groups 1, 2, and 3 under condition NI as contrasted with the relatively flat functions for the corresponding groups under the RI condition.

Figure 1 also shows that the last 60 trials provided the closest approximation to the asymptotic behavior required to assess the theories under consideration and Table 1 gives the proportions of A_1 responses for each group and condition for the last 60 trials. An analysis of variance of these response proportions showed a significant main effect of groups ($F = 3.21$, $df = 2/90$, $p < .05$) and a significant conditions by groups interaction ($F = 5.90$, $df = 2/90$, $p < .01$). Table 1 clarifies the interaction in that there appears to be a linear relationship among groups for condition NI whereas a quadratic

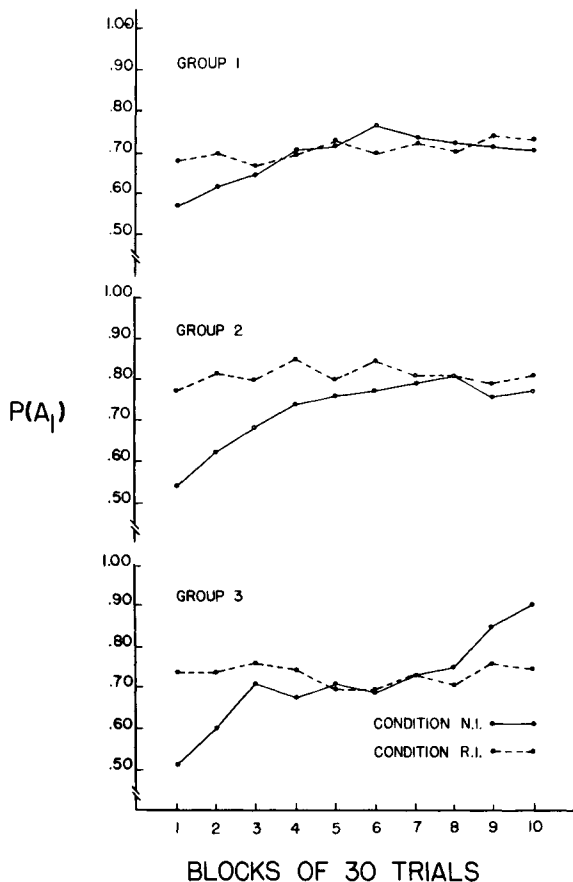


Fig. 1. Proportion of A_1 responses as a function of blocks of 30 trials for each group and condition.

Table 1. Proportion of A_1 Responses for the Last 60 Trials for Each Group and Condition

Condition	Group		
	1	2	3
N I	.713	.770	.875
R I	.755	.810	.765

function appears to describe the groups under condition RI.

The marginal response probabilities for the groups receiving the NI condition can be explained in terms of a utility for variability concept since $P(A_1)$ increased as a function of event variability. There is a clear tendency toward maximizing from a starting point close to probability matching for the group receiving only a single reinforcing event to extreme overshooting for Group 3. In fact, Fig. 1 suggests that this third group did not reach asymptote, and the relationship between the last two trial blocks indicates that the true asymptote may be somewhat higher than .875. The data obtained for the three groups under condition RI are more difficult to interpret. If the RI condition could be considered as providing greater amounts of variability than the NI condition (cf. Erickson, 1966), then a theory employing the concept of utility for variability would predict larger marginal response probabilities for each of the RI groups than for the corresponding NI groups. While Groups 1 and 2 provided the expected result in terms of the absolute values within conditions as well as the relative values between conditions, the data for the third RI group were not expected.

In general, Siegel's (1964) conception of utility for variability provided a good description of the marginal response probabilities for the NI condition which is similar to most probability learning tasks in that Ss learn the reinforcing event probabilities. The utility concept in general may prove useful for describing the data of several other experimental situations previously mentioned where probability overshooting is typically obtained.

References

- ERICKSON, J. R. On learning several simultaneous probability learning. *J. exp. Psychol.*, 1966, 72, 182-189.
- MESSICK, D. M. The utility of variability in probability learning. *Psychon. Sci.*, 1965, 3, 355-356.
- MESSICK, D. M., & RAPOPORT, A. A comparison of two payoff functions on multiple choice decision behavior. *J. exp. Psychol.*, 1965a, 69, 75-83.
- MESSICK, D. M., & RAPOPORT, A. Expected value and response uncertainty in multiple choice decision behavior. *J. exp. Psychol.*, 1965b, 70, 224-230.
- SIEGEL, S. *Choice, strategy and utility*. New York: McGraw-Hill, 1964.
- YOUNG, F. W. Twelve choice probability learning with payoffs. *Psychon. Sci.*, 1967, 7, 353-354.