Event memory in probability learning¹

ARTHUR S. REBER AND RICHARD B. MILLWARD

BROWN UNIVERSITY

Abstract

On each of 1600 trials of a probability learning experiment, Ss not only predicted the next event but also recalled the event on the x^{th} past trial ($x=1, \dots, 9$). The frequency distribution of x , $f(x)$, was the main independent variable. Ss did not match event probability but seemed to overshoot as a function of $f(x)$.

Problem

The study directly investigated the memory for the sequence of events in probability learning (PL). In two-choice PL with non-contingent reinforcement, Ss' response proportions generally match event proportions for a few hundred trials. Two broad classes of theories have been suggested to account for this result. One class involves complex dependence on the sequence of events, while the second assumes independence of sequence of events. Consider, as an example of the latter, statistical learning theory (SLT). It is assumed in SLT that each event increases, according to a linear operator, the probability of predicting that event. The theory assumes that the effects of events can be completely represented in the probability of responding; this is the independence-of-path assumption. An example of the former class of theories is Feldman's "cognitive" approach (1963). He assumes that Ss test hypotheses which are formed primarily on the basis of event or event-response sequences. The independence-of-path issue can also be stated in terms of memory. While there is no requirement to remember previous events in SLT, memory plays a very important role in the cognitive theories.

There is, of course, evidence to support either position. Friedman et al (1964) have reported PL data which can be very accurately described by SLT. Estes (1964) has cogently argued that many of the effects which seem contradictory to SLT (e.g., negative recency, peculiar variations in the response sequences) are simply temporary response biases which disappear with extended training. He assumes that there is a basic learning process very similar to that described by SLT which then may have higher-level response strategies imposed upon it. On the other hand, the evidence of experiments by Engler (1958), Anderson (1960), and Millward (1959) indicates that the sequence of events can and does act as a stimulus and significantly affects Ss' predictions.

Before attempting further work on sorting out the manner in which the sequences affect choice behavior, it seems reasonable to investigate the extent of the memory for sequences of events. In this study we took a very direct approach. We simply asked Ss to remember past events. If the results are to be useful for PL, the effects of such direct memory probes must be ascertained.

Method

The apparatus was a simple panel with an in-line digital projector, a response lever which could be thrown left or right, and two corresponding event lights. A trial consisted of two parts, a memory response and a prediction response. Each trial began with the presentation of a digit from 1-9 on a red background. The red background indicated to S that he was to recall an event on a past trial. The number indicated which past trial. For example, a "5" on a red background was a cue for S to remember which event occurred 5 trials back. Following S's memory response, the digit "1" appeared on a green background. This was a cue for S to predict the event that he thought would occur on that trial.

Three groups were run in which the distribution of the digits on memory trials differed. Let x ($x=1, 2, \dots, 9$) be a random variable defining the past trial for which S is to recall the event. Let $m(x)$ be the density function for x . The High group (H) had a density function $m(x) = x/45$; for the Rectangular group (R), $m(x) = 1/9$; and for the Low group (L), $m(x) = (10-x)/45$. A control group with no memory task (i.e., probability learning alone [PLA]) was also run. The groups had 8, 7, 7 and 7 Ss respectively. The Ss in the H, R and L groups were high-school seniors from the top half of their class; the Ss in the PLA group were undergraduates at Brown University.

Subgroups of 3 or 4 Ss each with different sequences of events were run for 8 days with 200 trials a day. On days 1 and 2, the probability of event E_1 (π) was .7; on days 3 and 4, $\pi = .1$; on days 5 and 6, $\pi = .6$; and on days 7 and 8, $\pi = .5$. On all days, all Ss in a given group had the same density function defining the memory trials.

Results

Figure 1 shows the proportion of A_1 prediction responses under each π condition for the four groups. The most obvious aspect of these results is that there is a great deal of "overshooting" and therefore the probability matching law does not hold. Note that in the H group the overshooting is greatest and somewhat greater than other researchers have noted. The amount of overshooting may be related to the length of the sequence "recalled" by the Ss. The H group had to recall an event of a distant trial ($x > 5$) more often than did the other two groups.

Figure 2 presents the proportion of correct memory recalls (P_C) as a function of the x value. All trials are combined regardless of the π value involved. The

Fig. 1. Proportion of A_1 prediction responses for each block of 50 trials.

Fig. 2. Proportion of correctly recalled probe trials over the entire experiment as a function of x :

H group does poorly on probe trials 1-5, indicating that emphasizing recall more than 5 trials back adversely affects recall performance on more recent trials.

Figure 3 shows the effect of the π value on memory for the past events. The form of the memory curve within each π value appears to be similar to the average over all π values. It declines rapidly through $x=6$ and then begins to level off. There seems to be an improvement in memory over days since P_C is about the same for $\pi = .7$ and $\pi = .5$, even though there is an obvious relationship between P_C and π . Because we ran only one order of π values, it is impossible to separate improvement and π value effects exactly.

Discussion

The results immediately raise several questions. For example, why should the H group predict the more frequent event more often? Three explanations have occurred to us. The first involves the difficulty of the task and the tendency for the Ss to "give up." The second, not unrelated to the first, is the possibility that Ss concentrate on the memory task and make a simple stereotyped response for the prediction task. The third possibility is that the random nature of the event sequence becomes obvious to Ss who pay attention to a longer-length sequence. Awareness that the event sequence is random is known to produce overshooting (Rubinstein, 1959). We cannot say at the moment which, if any, of these explanations is valid.

A second question concerns the tendency for recall performance to level off at 6 trials back. It certainly would be valuable to know whether 6 trials is the approximate limit of the memory for event sequences in the usual PL experiment. However, establishing the point at which a S's memory performance is no better than chance is difficult because we do not know how he guesses when he does not remember the event. The major problem lies in the complex interactions across π values between the probability that the S remembers correctly and the probability that he guesses correctly. For example, it is easier to be correct when $\pi = .1$ than when $\pi = .5$, but it is not clear how much of the increased probability of being correct can be attributed to ease of recall and how much can be attributed to

a change in the guessing rate. We do not know what a S's guessing rate is. An S might be guessing with a .5 probability, or matching the π value, or guessing at the same rate at which he predicts, or maximizing. A further possibility is that the method of encoding the event sequence for memory varies with the π value.

In summary, it is perhaps worth emphasizing that the prediction data of the R and L groups do not differ from the prediction data of the PLA group. This provides some support for our implicit assumption that asking about the previous event sequence does not distort the usual PL results. It also supports the notion, first presented by Goodnow & Pettigrew (1955) and emphasized by Anderson (1960) and Restle (1961), that Ss pay attention to the recent sequence of events. It does appear, however, that the H group was predicting differently from a standard PL procedure.

Fig. 3. Proportion of correctly recalled probe trials for each π value as a function of x .

References

- Anderson, N. H. Effect of first-order probability in a two-choice learning situation. *J. exp. Psychol.*, 1960, 59, 73-93.
- Engler, J. Marginal and conditional stimulus and response probabilities in verbal conditioning. *J. exp. Psychol.*, 1958, 55, 303-317.
- Estes, W. K. Probability learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 89-128.
- Feldman, J. Simulation of behavior in the binary choice experiment. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought*. McGraw-Hill, 1963. Pp. 329-346.
- Friedman, M. P., Burke, C. J., Cole, M., Estes, W. K., Keller, L., & Millward, R. B. Two-choice behavior under extended training with shifting probabilities of reinforcement. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Stanford University Press, 1963. Pp. 250-316.
- Goodnow, J. J., & Pettigrew, T. F. Effect of prior patterns of experience upon strategies and learning sets. *J. exp. Psychol.*, 1955, 49, 381-389.
- Millward, R. B. The discriminative aspects of stimulation from the reinforcers in a verbal conditioning experiment. M. A. Thesis, 1959. Indiana University.
- Restle, F. *The psychology of judgment and choice*. New York: Wiley, 1961.
- Rubinstein, I. Some factors in probability matching. *J. exp. Psychol.*, 1959, 57, 413-416.

Note

1. This research was supported in part by research grant G-18110 from the National Science Foundation, and by Brown University National Science Foundation Grant GU-448. The authors wish to thank Miss Edna Macdonald of Hope High School for her help in selecting subjects.