

Subjective probability revision in twelve simultaneous sequences¹

STANLEY G. LIPKIN and DARRELL D. WILLIS, Carnegie-Mellon University, Pittsburgh, Pa. 15213

Two groups of Ss gave successive probability estimates of coin bias in the same 12 sequences. The control group operated on sequences consecutively while the experimental group worked with all 12 sequences concurrently. The experimental group extracted more of the available certainty as measured by Bayesian predictions.

Several investigators have shown that subjective probability revision proceeds more gradually than would be justified by a Bayesian decision process (e.g., Phillips & Edwards, 1966; Peterson & DuCharme, 1967). For example, a S may be told that a given coin will show a .60:.40 bias over a long series of tosses. After observing each toss of the coin, the S is asked to guess the direction of the bias and to state his confidence in terms of odds or probabilities. Bayes' theorem provides the following model for revising the odds in favor of a bias toward heads (H_h) relative to tails (H_t) given the relevant datum (D) has occurred.

$$\frac{P(H_h | D)}{P(H_t | D)} = \frac{P(D | H_h)}{P(D | H_t)} \cdot \frac{P(H_h)}{P(H_t)} \quad (1)$$

or, posterior odds = likelihood ratio · prior odds. (See Edwards, Lindman, & Savage, 1963, for a detailed discussion of Bayes' theorem, or see Peterson, Schneider, & Miller, 1965, for an experimental example.) While the Bayesian model uses true probabilities to reach the optimal revision in posterior odds, human Ss typically make more conservative revisions.

It is hypothesized that one reason humans extract less certainty from data than does the Bayesian model is that they spend some of their information processing capacity operating on the specific sequential history. Bayesian revision, on the other hand, is a function of the difference between the cumulative number of heads and tails but not of the particular permutation of events comprising D (see Phillips & Edwards, 1966, p. 348, for further discussion). Thus, the purpose of the present study was to test whether subjective probability revision would be more nearly Bayesian under experimental conditions which mask the particular permutation of a sequence while preserving its cumulative history. The present technique for masking sequences was to have Ss work on 12 sequences simultaneously thereby making it extremely difficult to incorporate sequential information into their probability revisions.

METHOD

Subjects

Twenty male Carnegie-Mellon University students were randomly divided between the two experimental groups. Ss were paid \$1.25 per h for their participation.

Procedure

Each S was shown a stack of 120 3 x 5 cards which he was told recorded the results of a sequence of 10 tosses for each of 12 coins. With a brief demonstration, S was told that each of the 12 coins had been randomly selected from a pair of coins with .60:.40 biases in opposite directions. For both groups each 3 x 5 card showed (a) a Roman numeral from I to XII denoting the particular coin, (b) results of the latest toss, and (c) the cumulative number of heads and tails for the coin through the last toss. On this basis, S was instructed to guess the direction of the bias and to state his confidence rating in

terms of a probability from 0 to 100 after each of the 120 cards.

For Ss in the Successive Group, the cards were arranged so that the first 10 cards showed the 10 outcomes of flipping the first randomly chosen coin, the next 10 cards showed the outcomes from the second coin, etc. For Ss in the Simultaneous Group, the results of the first toss of each of the 12 coins was shown on the first 12 cards, the outcomes from the second toss of each coin on the second 12 cards, etc.

The sequences were made up from a table of random numbers. Of the 12 sequences, six were mirror images of the other six to control for heads-tails biases. Of the six unique sequences, two had eventual .6:.4 biases, two had .7:.3 biases, and two had .8:.2 biases. Half the sequences favored heads and half, tails. All Ss in each group received the same random order of sequences.

Dependent Measures

Comparisons between groups were based on the absolute deviation of subjective probabilities from Bayesian probabilities and accuracy ratios. The absolute deviations were preferred to the actual subjective probabilities since the prime issue of the study was to determine whether the Simultaneous Group was more Bayesian than the Successive Group. Accuracy ratios, defined as the subjective log likelihood ratio divided by the Bayesian log likelihood ratio, provide a trial-to-trial comparison with Bayesian revisions. (A full discussion of this measure is available in Peterson et al. 1965.)

RESULTS

Absolute Deviations from Bayes

The observed effect of having Ss revise probability estimates in 12 sequences concurrently rather than consecutively was a closer approximation to Bayesian predictions. With probability estimates averaged over mirror-image sequences and over Ss, there are 60 available comparisons between the Simultaneous and Successive groups, corresponding to 10 trials on each of the six unique sequences. In 53 of these comparisons, the Simultaneous group was closer to the Bayesian prediction. A

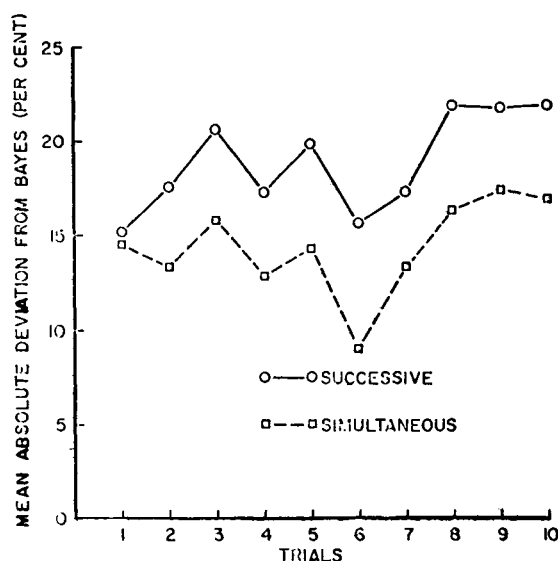


Fig. 1. Absolute deviations from Bayesian predictions averaged over sequences and Ss.

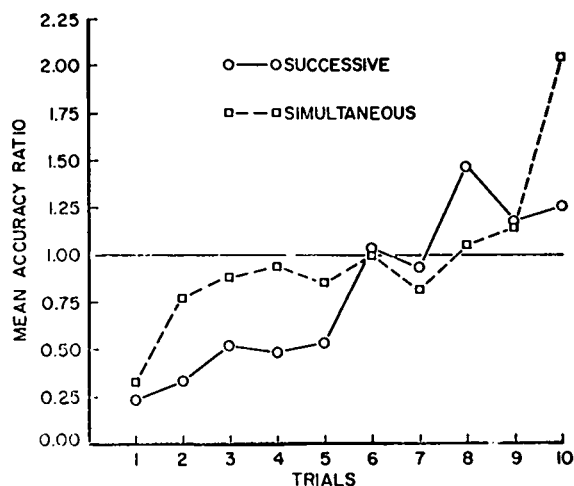


Fig. 2. Accuracy ratios as a function of trials averaged over sequences and Ss.

Wilcoxon matched-pairs signed-ranks test (Siegel, 1956) which uses both direction and amount of difference, clearly shows the Simultaneous group to be closer to Bayesian predictions, $z = -6.1$, $p < .001$. Figure 1 shows this effect as a function of trials averaged over sequences.

Virtually all probability estimates were less extreme than the Bayesian predictions. There were no reliable exceptions to this conservatism in the Successive group. In the Simultaneous group, on the other hand, there was one consistent exception to this general finding of conservatism. Ss in the Simultaneous group almost unanimously matched the Bayesian prediction of no bias whenever there was no difference in the cumulative number of heads and tails.

Accuracy Ratios

The present experimental manipulation also appears to have affected the amount of trial-to-trial revision relative to Bayesian predictions. In Fig. 2, the Bayesian prediction is indicated by the horizontal line. The functions for the Simultaneous and Successive groups both have positive slope indicating that Ss tended to under-revise relative to Bayes early and to over-revise later. The two functions appear to have similar slopes, and analysis of variance indicated no significant difference between the groups, $F < 1.0$. However, a closer comparison between the two groups suggests that the Simultaneous group is substantially Bayesian between Trials 2 and 9 with little indication of positive slope, while such does not appear to be the case for the Successive group. To test the reliability of this tendency, the groups were compared on the basis of the absolute deviations of accuracy ratios from 1.0. Using the average accuracy ratios from the 60 sequence-trials, a Wilcoxon matched-pairs signed-ranks test showed that the Simultaneous group was generally closer to the Bayesian revision, $z = -3.99$, $p < .001$. In 48 of the 60 comparisons, the Simultaneous group's average accuracy ratio was closer to 1.0 indicating that the average function shown in Fig. 2 is representative of individual sequences.

DISCUSSION

The principle result is that Ss who revised probability estimates on 12 sequences at once were able to extract more of the available certainty than Ss who dealt with each sequence individually. This finding may be understood in terms of different information processing strategies. Since

sequences were presented intact to Ss in the Successive Group, it would have been relatively easy for these Ss to have retained the order of event occurrences and possibly to have used this information in probability revision. This speculation is based on the common finding that Ss tend to predict the occurrence of randomly scheduled events as if there were sequential dependencies present (e.g., Jarvik, 1951; Feldman, 1962). Such a strategy is obviously nonoptimal. However, any distracting search for sequential dependencies would have been an overwhelming task for Ss in the Simultaneous Group. Thus, the closer approximation to Bayesian predictions in the Simultaneous Group may be attributed to a forced strategy of basing probability estimates more heavily on cumulative event counts.

The present results are similar to those of probability learning experiments which have employed methodology affecting the use of sequential information. Peterson & Ulehla (1965) discouraged search for sequential dependencies by having one group of Ss generate the random sequences of events themselves. These Ss showed a significantly greater proportion of maximizing responses than a control group who responded to the same sequences without prior knowledge of their generating source. Multiple-cue probability experiments afford an even closer parallel to the present study since both call for probabilistic decisions in several concurrent sequences. Erickson (1966) found that choice proportions in a four-game situation showed more maximizing than is typically observed in a single probability learning game. Similarly, Peterson, Hammond, & Summers (1965) found nearly optimal responding as measured by a linear multiple-regression equation in a three-game probability learning situation. The general implication from these findings and those of the present experiment is that one cause of conservatism is the tendency of Ss to employ sequentially dependent strategies in tasks with random sequences. A procedure which inhibits the use of sequential information while preserving the more elementary and necessary cumulative data leads to more optimizing.

REFERENCES

- EDWARDS, W., LINDMAN, H., & SAVAGE, L. J. Bayesian statistical inference for psychological research. *Psychological Review*, 1963, 70, 193-242.
- ERICKSON, J. R. On learning several simultaneous probability-learning problems. *Journal of Experimental Psychology*, 1966, 72, 182-189.
- FELDMAN, J. Computer simulation of cognitive processes. In H. Borko (Ed.), *Computer application in the behavioral sciences*. Englewood Cliffs, N. J.: Prentice-Hall, 1962. Pp. 336-359.
- JARVIK, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, 1951, 41, 291-297.
- PETERSON, C. R., & DuCHARME, W. M. A primacy effect in subjective probability revision. *Journal of Experimental Psychology*, 1967, 73, 61-65.
- PETERSON, C. R., HAMMOND, K. R., & SUMMERS, D. A. Optimal responding in multiple-cue probability learning. *Journal of Experimental Psychology*, 1965, 70, 270-276.
- PETERSON, C. R., SCHNEIDER, R. J., & MILLER, A. J. Sample size and the revision of subjective probabilities. *Journal of Experimental Psychology*, 1965, 69, 522-527.
- PETERSON, C. R., & ULEHLA, Z. J. Sequential patterns and maximizing. *Journal of Experimental Psychology*, 1965, 69, 1-4.
- PHILLIPS, L., & EDWARDS, W. Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 1966, 72, 346-354.
- SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

NOTE

1. This study was supported by NIH Grant MH 07722. Thanks are due to Karen S. Wiseman for help in analyzing the data and to Vincent E. Bitetto for his comments.