Response probabilities in a verbal discrimination task with probabilistic reinforcement[1]

IRWIN P. LEVIN and J. FRANK DOOLEY, University of Iowa, Iowa City, Iowa 52240

Forty Ss received 160 presentations of each of four word-pairs in a variation of the noncontingent binary-choice probability learning paradigm. Successive presentations of any given word-pair were separated by the interpolated presentations of other word-pairs. For each pair, one of the words was designated as correct on a randomly-chosen 70% of the presentations while the other word was correct on the remaining 30% of the presentations. Inspection of asymptotic response probabilities, which approximated the reinforcement probabilities, and first-order conditional probabilities indicated that this method of stimulus presentation may provide a less ambiguous test of the assumptions of certain mathematical models than is afforded by the standard probability learning paradigm.

Criticism of the application of relatively simple mathematical models (e.g., the model of Estes & Straughan, 1954) to the noncontingent binary-choice probability learning paradigm rests primarily on the inability of the models to take into account the effect of the preceding sequence of reinforcing events and responses, that is, the nonindependence of individual trials (cf. Anderson 1960, 1964, 1966). The probabilistic verbal discrimination paradigm employed in the present study represents an attempt to reduce the effect of "trace stimuli" upon the S's response by interpolating between successive presentations of any given word-pair the presentation of other word-pairs. The list of word-pair presentations was so constructed that one to six such interpolations were presented before reoccurrence of a given word-pair. One member of each pair was designated "correct" on a randomly chosen 70% of the presentations of that pair, while the other member was designated "correct" on the remaining 30% of the presentations.

The present paper demonstrates the correspondence between the results with this paradigm and the gross behavioral phenomena found in the probability learning situation. Asymptotic response probabilities and first-order conditional probabilities are computed for the presentations of individual pairs to test the correspondence to the usual paradigm with respect to the typical "matching" behavior (cf. Estes, 1961) and the sensitivity of S's response to the effect of the immediately preceding presentation of the same word-pair.

The current study represents an attempt to develop a paradigm which might reduce or eliminate the action of memory traces of preceding reinforcement events and responses which may contaminate the operation of the associationistic process described by Estes' formulations.

METHOD
Four word-pairs consisting of high-frequency words with minimal within-pair or between-pair interassociations were selected. These pairs are listed in Fig. 1. The pairs were presented on a Stowe memory drum at a presentation interval of 2.5 sec. A list of 16 word-pairs (four presentations of each pair) was so constructed that one to six interpolations occurred between each successive presentation of a given pair. This list was repeated until each individual word-pair had been presented 160 times. The left-right position of each member of a pair was counterbalanced over trials, and Ss were instructed that the position of a word within a pair did not determine the "correctness" of that word.

A 70-30 reinforcement schedule was maintained within each block of 40 trials; that is, within each block of 10 presentations of a given word-pair one member of the pair was designated as correct on a randomly-chosen seven of the trials and the other member was designated as correct on the remaining three trials. The S was instructed to try to anticipate the "correct" word on each presentation. Following S's oral response, E designated the correct response orally.

Forty male students from undergraduate psychology courses at the University of Iowa were participants in the experiment.

RESULTS
Figure 1 shows the probability of the 70%-reinforced response, P(A1), plotted over successive blocks of 20 presentations of each of the four word-pairs. The word listed first in each pair is the one reinforced 70% of the time for all Ss.

Probability on Trial 0 represents the relative frequency of A1 given before the first reinforcement. The high initial value for the pair, light-ball, the first pair presented to each S, may be due to an initial response bias based upon a combination of word preference and right-left position bias in reading the words. Further experimentation has shown that counterbalancing of word position on the first presentation tends to reduce this bias, but that counterbalancing across Ss of the 70%-reinforced word within each pair is necessary to eliminate bias. However, inspection of Fig. 1 reveals that response rates are relatively homogeneous between pairs after the first block of trials, which would seem to indicate that asymptotic response probabilities tend not to be affected by initial response biases.

For three of the four pairs, the asymptotic response probabilities, P(A1) averaged over the last two blocks of trials, were slightly greater than the "matching" level of .70; for the remaining pair the asymptotic response probability was slightly less than the "matching" level. Averaged over pairs and averaged over the last two blocks of trials, the mean value of P(A1) was .726.

These mean values do not necessarily reflect the tendencies of individual Ss to "match" the reinforcement contingencies or to "maximize" their expected number of correct responses. To give an index of the individual-difference factor, the relative frequency distribution of values of P(A1), averaged over the last two blocks of presentations of each pair, was computed. P(A1) was found to be .90 or greater for 8.8% of the S-item combinations; P(A1) was between .80 and .89 for 27.5% of the S-items; P(A1) was between .70 and .79 for 23.1% of the S-items; P(A1) was between .60 and .69 for 18.1% of the S-items; P(A1) was between .50 and .59 for 12.5% of the S-items; and P(A1) was below .50 for the remaining 10.0% of the S-items. Thus, in very few cases were less than half A1 responses made and in very few cases did Ss tend to "maximize" reinforcement by making almost all A1 responses. For most S-item combinations, the asymptotic response probability was within .20 of the matching level of .70. The number of Ss who responded at a rate above matching was somewhat greater than the number of Ss who responded at a rate below matching.

The aforementioned data summaries describe how Ss' response rates tend to correspond to the overall reinforcement probability. It is also of interest to see how Ss' behavior is influenced by the short-term response and reinforcement contingencies. Specifically, the following conditional proba-
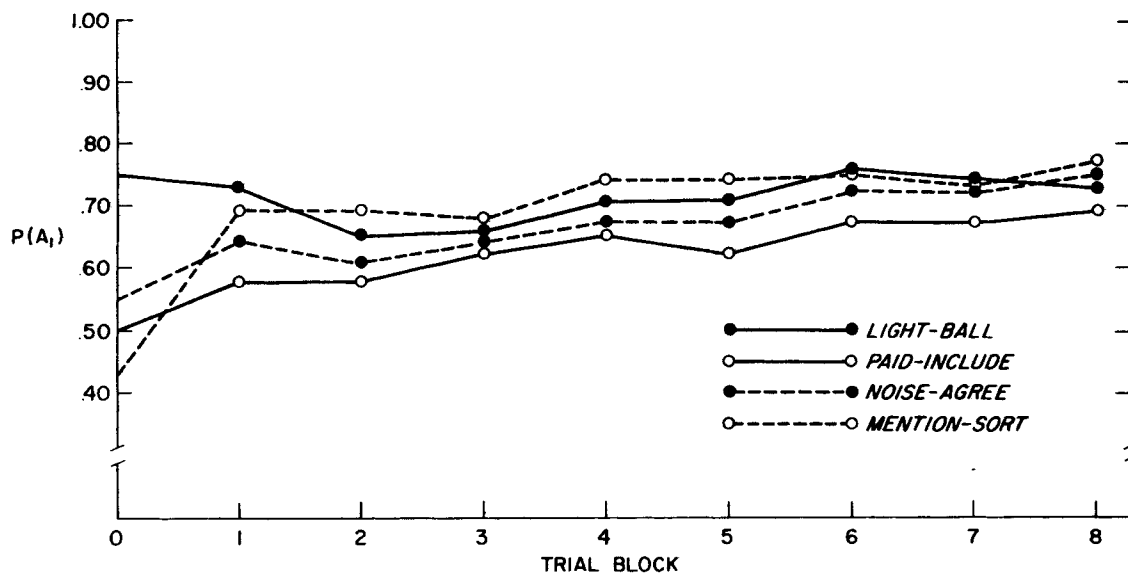
Fig. 1. Probability of the 70%-reinforced response, P(A1), for successive blocks of 20 presentations of each of the four word-pairs.

bilities were computed for each S: The probability of an A1 response following the presentation of a specific pair, given that an A1 response had been made on the previous presentation of that pair, $P(A1/A1)$; the probability of an A1 response given that the opposite response (A2) had been made on the previous presentation of the pair, $P(A1/A2)$; the probability of an A1 response given that on the previous presentation the A1 response had been indicated as "correct," regardless of the response actually made, $P(A1/E1)$; the probability of an A1 response given that the opposite response had been indicated as "correct" on the previous presentation, $P(A1/E2)$.

The mean conditional probabilities, averaged over the last two blocks of trials, with the appropriate standard deviations reflecting the magnitude of individual differences, are given for each word-pair:

light-ball: $P(A1/A1) = .734$ $(s = .187)$; $P(A1/A2) = .713$ $(s = .210)$; $P(A1/E1) = .744$ $(s = .173)$; $P(A1/E2) = .71$ $(s = .173)$

paid-include: $P(A1/A1) = .663$ $(s = .170)$; $P(A1/A2) = .706$ $(s = .177)$; $P(A1/E1) = .688$ $(s = .160)$; $P(A1/E2) = .665$ $(s = .158)$

noise-agree: $P(A1/A1) = .738$ $(s = .184)$; $P(A1/A2) = .692$ $(s = .220)$; $P(A1/E1) = .754$ $(s = .191)$; $P(A1/E2) = .691$ $(s = .165)$

mention-sort: $P(A1/A1) = .736$ $(s = .165)$; $P(A1/A2) = .767$ $(s = .203)$; $P(A1/E1) = .745$ $(s = .144)$; $P(A1/E2) = .775$ $(s = .147)$

It can be seen that all of these conditional probabilities correspond within .05 of the asymptotic response probabilities seen in Fig. 1. That is, on the whole, any effects attributable to the particular response or reinforcement given on the previous presentation of a given pair are relatively small in magnitude.

DISCUSSION

The use of the current paradigm seemed to be successful in reducing the contingency of the response to a given pair upon the previous reinforcement or response given to that pair. The asymptotic response levels observed with the current paradigm

approximated the "matching" behavior usually observed in the standard probability learning task. However, in the present study, there was a slight tendency to "overshoot" the matching level, with the asymptotic response probability being greater than .70 for three of the four pairs. Overall matching behavior has been observed in probability learning situations even though negative recency effects and response alternation effects were observed over short sequences of trials (Anderson & Whalen, 1960). In the present situation, the elimination of these effects could have served to raise the mean asymptotic value of P(A1).

Before a detailed theoretical analysis of behavior in the probabilistic verbal discrimination task can be made, it would seem wise to tabulate the empirical findings over a wide range of reinforcement probabilities and with varying numbers of items in the list. In the latter regard, it is of particular interest to investigate how the number of interpolated items affects Ss' tendency to respond on the basis of short-term patterns of previous responses and reinforcements. This is the subject of current work.

REFERENCES

ANDERSON, N. H. Effect of first-order conditional probability in a two-choice learning situation. Journal of Experimental Psychology, 1960, 59, 73-93.

ANDERSON, N. H. An evaluation of stimulus sampling theory: Comments on Professor Estes' paper. In A. W. Melton (Ed.), Categories of human learning. New York: Academic Press, 1964. Pp. 129-144.

ANDERSON, N. H. Test of prediction of stimulus sampling theory in probability learning. Journal of Experimental Psychology, 1966, 71, 499-510.

ANDERSON, N. H., & WHALEN, R. E. Likelihood judgments and sequential effects in a two-choice probability learning situation. Journal of Experimental Psychology, 1960, 60, 111-120.

ESTES, W. K. A descriptive approach to the dynamics of choice behavior. Behavioral Science, 1961, 6, 177-184.

ESTES, W. K., & STRAUGHAN, J. H. Analysis of a verbal conditioning situation in terms of a statistical learning theory. Journal of Experimental Psychology, 1954, 47, 225-234.

NOTE