Test of a duoprocess paired-associate learning model in a simultaneous presentation situation

IRWIN D. NAHINSKY
UNIVERSITY OF MISSOURI

The predictive powers of two all-or-none PA learning models were tested. in a simultaneous presentation situation. A duoprocess model which assumed all-or-none elimination of errors proved superior to the model which did not postulate error elimination. A tendency to match each alternative to one stimulus was found.

Nahinsky (1964) has developed an extension of the one-element stimulus-sampling model for paired-associate (PA) learning (Bower, 1961). This duoprocess (DP) model postulates a one stage elimination of incorrect Rs in addition to the one stage learning of correct Rs assumed by the uniprocess (UP) Bower model.

Most studies of PA learning which involve tests of mathematical models use the successive presentation method. The present experiment involves a simultaneous presentation technique with fixed order of trials which maximizes possible interactions among stimuli during learning. Both positive and negative feedback conditions were used. Because both models postulate that single S-R associations are learned independently of one another, departures from predictions of the models in this simultaneous presentation situation may help us to gain insight into the nature of interdependencies among stimuli in a list.

A brief summary of assumptions of the DP model is in order. Consider the PA situation in which S must learn a correct R from k possible alternatives for each of a number of stimuli. The model assumes:

1. The probability of a correct R for each stimulus is 1/k initially with probability c of being learned upon occurrence after which it always occurs to the stimulus.

2. The probability of an incorrect R for a given stimulus is (k-1)/k initially, with probability c of being eliminated permanently upon occurrence, after which the correct R probability jumps to 1/(k-1). There are corresponding jumps in correct R probability with each successive incorrect R elimination.

Experiment

Two groups of Ss learned Rs to a set of eight CVC trigrams. The trigrams were selected randomly from a list having 20% association value as determined by Glaze (1928). The vowel E appeared in three syllables, A in two syllables, and the other three vowels in one syllable each. No syllable had more than one letter in common with any other syllable. A pasteboard screen separated E and S from each other's view. The learning criterion for each group was two consecutive cycles of correct Rs for all stimuli.

Group I (positive feedback condition): Each S in this group, consisting of 18 university students, was presented with a sheet containing all of the syllables in the same specified random order on each trial. S was told that the task involved finding the right number between 1 and 8 for each stimulus and placing that number in the blank next to the syllable. He was told that each number could be used as an R to any number of stimuli or to no stimuli. Each S was instructed specifically to learn the correct associations as quickly as possible. After S filled out a sheet, E marked the correct Rs "OK" and marked nothing for other Rs. S was allowed to inspect his corrected paper for 20 sec. between trials.

Group II (negative feedback condition): This group consisted of 16 university students. The conditions duplicated those for Group I, except that Ss were instructed specifically to eliminate wrong Rs as well as to learn correct Rs. Here E marked wrong Rs with a red check and marked nothing for correct Rs.

Results and Discussion

Estimates of parameters outlined in the models were made from the data for both conditions. The values of c for both models were estimated from the average number of errors over Ss in each condition (see Bower, 1961; Nahinsky, 1964). An empirical estimate of c relevant to the theoretical comparison can be derived from the all-or-none assumption common to both theories. This assumption suggests that we consider an R conditioned for a stimulus on the trial after the final error and one subsequent correct R. Since correct responding invariably follows the correct R, we might assume the R is conditioned on that trial. By taking the ratio of associations to be learned to total correct Rs during these unlearned association trials, we get an estimate of the probability an association is made when the correct R occurs, i.e., an estimate of c. This may provide a slight overestimate of c, because of chance correct responding after the estimated learning point. Nevertheless, the estimate illustrates some points of interest.

Table 1. Comparison of Uniprocess and Duoprocess Models for Predictive Power of Parameter Estimates (N=34)

Variable	Empirical estimate	Uniprocess prediction	Duoprocess prediction
c	.687	1.240	.620
Total proportion of correct Rs to total Rs for unlearned association trials	.205	.125	.222
Total number of correct Rs for unlearned association trials	M=11.65 SD_M=1.20	7.10	12.90

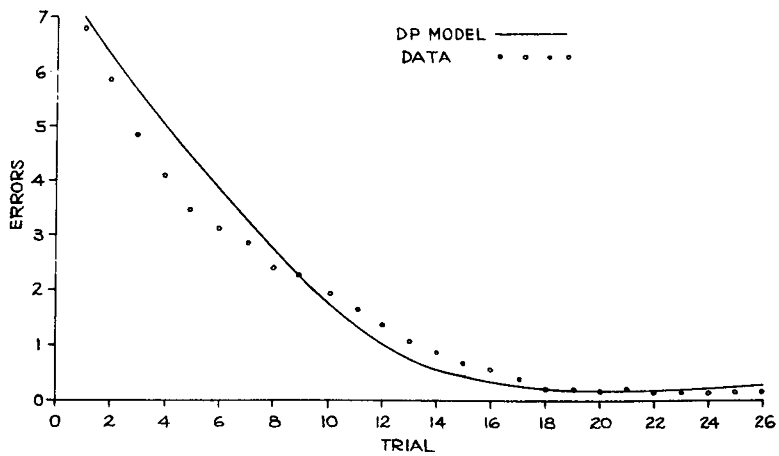

Fig. 1. Simultaneous presentation data: observed and predicted error curves for DP model. (N = 34)

The positive feedback procedure yielded an empirical c estimate of .681 and the negative feedback procedure a value of .692. These two procedures resembled each other so closely with respect to other values of interest that data for these two conditions was combined. Table 1 shows comparisons between observed values and the two models. The predicted DP c was .62, the corresponding UP value 1.24, and the empirical estimate .687. Since c is a hypothesized probability, it is difficult to interpret the UP c. Therefore, the empirical c estimate was used in other UP predictions. These data obviously favor a DP interpretation.

Estimated total correct Rs during unlearned association trials and proportion of correct Rs to total Rs for unlearned associations also show a better fit for DP than UP assumptions. After computing the observed mean and variance for total correct Rs during unlearned association trials, a t test was used to test the null hypothesis that the observed mean was drawn from a population in which the true mean equalled the theoretical mean for each model. For the UP model, $t = 3.73$, $df = 33$, $p < .01$, and for the DP model, $t = 1.02$, $df = 33$, $p > .05$. Thus, results show that only the UP model deviated significantly from the data.

The DP model generates a unique prediction about stationarity of correct R probabilities before the final error. Consider trials until the first trial upon which some wrong alternative occurs for the last time to some stimulus. There should be stationarity of correct R probabilities during these trials, because no transition can occur until some previously used incorrect alternative occurs for the last time. Conversely, it is also predicted that nonstationarity would obtain for trials after this estimated first transition trial until the final error. A X^2 test was performed for stationarity (see Suppes & Ginsberg, 1963) over blocks of three trials until the estimated first transition point. The resulting X^2 led to acceptance of the null hypothesis ($X^2 = 1.12$, $df = 2$, $p > .50$, total Rs = 418). These results give strong support to stationarity predicted by both models. The same analysis was performed for blocks of three trials after the estimated first transition trial until the final error. The observed value ($X^2 = 27.76$, $df = 5$, $p < .001$) also led to the rejection of the null hypothesis. Thus stationarity breaks down in trials after the possible elimination of a wrong alternative as predicted by DP assumptions but not by UP assumptions.

Figure 1 shows the observed and predicted error curves for the data. The corresponding UP predicted curve was not plotted because of the dubious c value. Although the curve appears to fit fairly well, there is a regular deviation which required explanation. The almost open invitation to matching in the simultaneous situation seems to provide the key. The tendency to use all eight alternatives on each trial for the eight stimuli was universal in initial phases despite instructions. Because only two alternatives were correct for more than one stimulus, this matching tends to produce faster initial progress than the model predicts. The observed error curve is lower than the predicted curve for the first eight trials which reflects this tendency. However, persistence in matching in later trials would slow progress, as is evidenced by a slower drop in errors than predicted between trials 9 and 17. Nonetheless the other evidence, particularly the stationarity evidence, indicates a DP interpretation is feasible. Additional complexities imposed by matching behavior would have to be incorporated into the model to account adequately for learning in such situations.

The matching tendency evident in simultaneous data should make us cautious about using the simplifying assumptions of independence among stimulus and response elements for PA tasks in general. Despite emphasis upon such independence in instructions to Ss there appears to be a strongly ingrained tendency to view the elements of the task as an interconnected whole.

References

Bower, G. H. Application of a model to paired-associate learning. *Psychometrika*, 1961, 26, 255-280.

Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, 35, 255-267.

Nahinsky, I. D. A duoprocess model for paired-associate learning. *Psychol. Rep.*, 1964, 14, 467-471.

Suppes, P., & Ginsberg, Rose. A fundamental property of all-or-none models, binomial distribution or responses prior to conditioning, with application to concept formation in children. *Psychol. Rev.*, 1963, 70, 139-161.