Recognition operating characteristics as a function of prior recall confidence[1]

BRUCE A. PAPPAS AND MILTON D. SUBOSKI

QUEEN'S UNIVERSITY, KINGSTON, ONTARIO

Ss rated their confidence in the correctness of their response in a recall-recognition RTT paired-associates paradigm. For items correct on T_1, T_2 recognition operating characteristics showed increasing recognition accuracy as a function of T_1 confidence. Items incorrect on recall produced an ROC indicating nonrandom performance on T_2. The results are interpreted as opposing an all-or-none theory of paired-associates learning.

A recent application of signal detection theory has suggested that recognition memory is a continuous rather than a state process (Murdock, 1965). Using confidence ratings in an RTRT paradigm, Suboski, Pappas, & Murray (1966a) have suggested that recall, as well, displays a continuum of states of associative strength. In a subsequent RTT study (Suboski, Pappas, & Murray, 1966b), the probability that an item would be recalled on T_2 was found to be a function of confidence rating on T_1. The present study uses a modified RTT paradigm, in which T_1 is a recall test, and T_2 a recognition test. The earlier results and the hypothesis that recall confidence is an indicator of associative strength predicts separate recognition operating characteristic (ROC) curves for recognition items as a function of the confidence rating of those items on the prior recall test. Relative positions of the curves would indicate the accuracy of recognition performance.

Method

A group of nine volunteer secondary school students were presented with ten lists, each consisting of a random pairing of the letters A-J as stimulus items and the numbers 0-9 as responses. The S-R pairs were placed on 3 x 5 in. blank cards and presented at a 4 sec. per pair exposure rate to all Ss by a manual "flash-card" technique. After a single list presentation, Ss worked, at the rate of 7 sec. per page, through a 10-page booklet. Each page contained a stimulus letter, a space to write in the response number, and a 5-cm "confidence" line, which had the left end labelled "guess" and the right end labelled "sure." The Ss were instructed to draw a vertical line across the "confidence" line to indicate their degree of confidence in the correctness of their responses. The order of items within booklets was constant across Ss and random across tests and lists.

After the recall test, Ss worked on an association task for 1 min. They were informed that this was an important part of the experiment. Immediately after the association task, Ss were presented with a second booklet, similar to the first except that responses had already been written in. Half of the responses were

correct and half incorrect. Each booklet had a different random order of correct and incorrect. The Ss' instructions on the recognition test were to score each pair as to correct or incorrect and again rate their confidence as for recall. They had been informed that some items would be incorrect but not the exact proportion. Responding was again at a 7 sec. rate.

Results

Confidence categories were successive 1-cm lengths of the confidence line measured from the left (guess) end. Proportion of responses correct on T_1 as a function of confidence rating, pooled across lists and Ss, were .188, .124, .337, .404 and .867 for categories one through five respectively. With the exception of the one inversion, the results are similar to the monotonically increasing functions obtained in previous studies (Suboski, Pappas, & Murray, 1966a, b).

The ROCs for T_2 were constructed by determining the proportion (pooled across Ss) of assigning confidence category i or stricter (Ri), to a recognition pair and cumulating from the strictest (yes, 5) to the most lax (no, 5) criterion (Egan, 1958). These conditional proportions were determined separately for those pairs which were correct (S-R), i.e., identical with the originally presented pairs, and those that were incorrect (S-$\bar{R}$). Figure 1 displays four ROCs, three as a function of rating for items correct on T_1 and the fourth
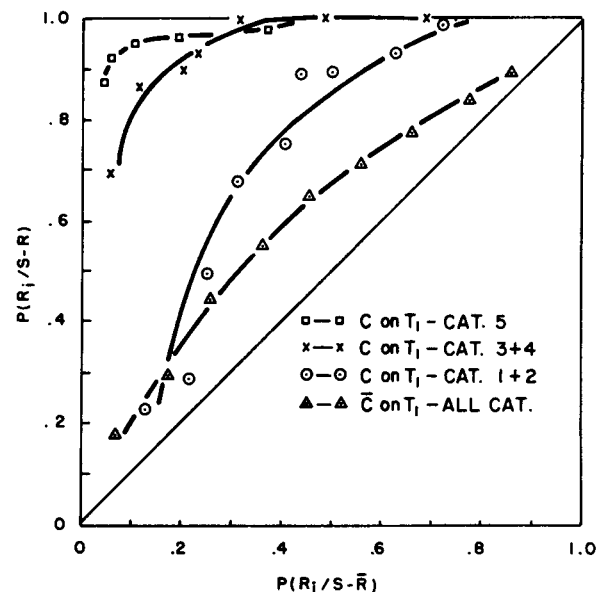

Fig. 1. T_2 ROCs as a function of correct or incorrect and rating categories on T_1.

for all items incorrect on T_1, constructed by plotting the conditional proportions of assignment of category i or stricter for S-R pairs on the ordinate and the corresponding conditional proportion for incorrect pairs on the abscissa.

The recognition accuracy represented by each curve is a function of the upward-left displacement of the ROC from the positive diagonal (chance line). The order relationship of the curves for items correct on T_1 indicates increased recognizeability as a function of T_1 ratings. The ROC for items incorrect on T_1 indicates accuracy somewhat less than that of low confidence corrects, but displacement from the chance line inconsistent with a purely random decision process. Separate ROCs for incorrect items as a function of T_1 confidence (not shown) failed to demonstrate any relationship between T_1 rating and T_2 ROC.

Discussion

The results show that recognizeability of correctly recalled items is a function of at least three separate states of recall confidence. A fourth state is suggested for items recalled incorrectly. However, since the confidence categories were arbitrarily selected, specifying the number of states is difficult. At least four are required and a continuum is consistent with the present results.

When recall is incorrect, information concerning the correct pairing is nevertheless evidenced on a subsequent recognition test. Correctly rejecting an incorrect pairing by having learned that the response is linked with another stimulus may partially account for the uppermost four points of the T_1 incorrect ROC. However, the "hits" to "false alarm" proportions which determine the lower five points are not directly affected by such information and yet indicate accuracy substantially above chance. These results are subject to several interpretations. Associative strength may vary for some items from the first to second test. Jones (1962) has shown that recall probabilities increase across repeated tests and confidence was also shown to increase across tests for N-C items in the RTT paradigm (Suboski, Pappas, & Murray, 1966b). Such changes could result from the application of information gained in the first test. The fact that incorrectly recalled responses are not reliably rated also suggests variation from test to test. The low associative strength of wrongly recalled items may randomly fluctuate from first to second test. A second possible explanation is that wrongly recalled items might have associative strengths which are too uniformly low to be rated reliably, but which nevertheless are sufficient to affect accuracy on the easier recognition task.

Interpretation of the data in terms of an all-or-none theory of recall such as that proposed by Estes (1960) does not seem feasible. Recall confidence ratings have previously been shown to reliably estimate the probability of correct recall (Suboski, Pappas, & Murray, 1966a, b), and in the present study to index subsequent recognition accuracy. These results agree with those of Williams (1962) based on recall response latencies, in suggesting multiple and possibly continuous states of recall associative strength.

References

Egan, J. P. Recognition memory and the operating characteristic. Indiana University Hearing and Communication Laboratory, Tech. Note AFCRC-TN-58-51, 1958.

Estes, W. K. Learning theory and the new "mental chemistry." *Psychol. Rev.*, 1960, 67, 207-223.

Jones, J. E. All-or-none versus incremental learning. *Psychol. Rev.*, 1962, 69, 156-160.

Murdock, B. B., Jr. Signal-detection theory and short-term memory. *J. exp. Psychol.*, 1965, 70, 443-447.

Suboski, M. D., Pappas, B. A., & Murray, D. J. Confidence ratings in recall paired-associates learning. *Psychon. Sci.*, 1966a, 5, 147-148.

Suboski, M. D., Pappas, B. A., & Murray, D. J. Confidence ratings in recall paired-associates: The RTT paradigm. *Psychon. Sci.*, 1966b, 5, 315-316.

Williams, J. P. A test of the all-or-none hypothesis for verbal learning. *J. exp. Psychol.*, 1962, 64, 158-165.