A direct investigation of hypothesis-making behavior in concept identification¹

RACHEL FALMAGNE, *University of Brussels,² Brussels, Belgium*

Two groups of Ss were faced with the same concept identification task. The first group performed under standard conditions, while the Ss from the second group had to give confidence ratings about the hypotheses they were entertaining. These ratings permitted a prediction of the first group's performance.

In an earlier paper (Falmagne, 1968) it had been supposed that the subject in a concept identification task was behaving according to hypotheses about the correct concept, and that the basic learning process was the modifications in the "strength" (interpreted either as a subjective likelihood or as the intensity of a memory trace) of these hypotheses.

Hypothesis-making behavior has often been assumed to be the basis of concept identification (e.g., Bower & Trabasso, 1963; Restle, 1962; Trabasso & Bower, 1966). Several instances can be found where that behavior had been studied directly by methods that made the current hypotheses of the subject observable (Levine, 1963, 1966; Richter, 1965; Suppes & Schlag-Rey, 1965); however, all these methods implied a distortion of the experimental situation with respect to the standard one.

The purpose of the present study was twofold: first, to provide an experimental technique which would render the hypothesis-making behavior, and more particularly, the modifications of the strengths of hypotheses, quasi-observable, without altering considerably the initial situation; second, to examine the relevance of the notion of hypothesis-making and of the variable "strength" in predicting concept identification performance.

The experimental technique used here consisted in recording in a standard task, trial-by-trial confidence judgments of the Ss about the various hypotheses they were entertaining. These confidence judgments were then used as estimates of the corresponding strengths to predict the behavior of another group working under the usual conditions.

METHOD

Stimulus Material

The stimulus set was composed of 64 different lines each of which was shown on a 2½ x 4 in. card. The lines could be either large (approximately 1½ in.) or small (approximately 1 in.), either horizontal or vertical, either single or double, either straight or curve, either at the top or at the bottom of the card, and they could be either dotted or plain lines.

Stimuli were classified by the E into As and Bs, according to whether they were plain or dotted, and the appropriate letter was written on the back of each stimulus card. The 64 stimulus cards were arranged in a deck on the table facing the S.

Subjects and Procedure

Two groups of 20 adult volunteers were run individually. The Ss were told that they would have to discover a systematic classification rule using the procedure described below. They were given detailed instructions about the description of the stimulus set and about the fact that the E's classification rule would be a "simple" one; that is, would be based upon a single characteristic of the stimuli. Several examples of the stimuli and the verbal statement of the 12 simple rules were shown. In the subsequent analysis, it was therefore assumed that the Ss never entertained hypotheses different from the ones corresponding to these 12 rules.

In both groups, on each trial, the S picked up a new stimulus card from the deck. Ss from the first group (Group R) then had to guess which class the stimulus belonged to (either A or B), and to state the answer verbally. They then turned the stimulus card over and examined the correct answer. In the second group (Group H), when a stimulus was presented, the S had to turn the card over and examine the right answer immediately. He then was

given a deck of 12 index cards each of which showed the statement of a simple rule. The S was instructed to write under each rule an integer ranging from 0 to 10 as an increasing function of his confidence in the rule being the E's rule. He was told to use 0 and 10 only in case of, respectively, negative and positive certitude.

In both groups as soon as the response and the reinforcement of a trial were over, the S picked up the next stimulus card from the deck. The task was self-paced; however, the Ss were instructed (and occasionally reminded) to work as fast as they could and to try to discover the rule in the shortest time and in the least number of trials possible.

The sequence of stimuli was the same for all Ss. Ss of Group R were run up to a criterion of 10 consecutive correct responses; Ss of Group H were run until they, on three consecutive trials, gave Score 10 to the correct rule and 0 to all the others.

RESULTS³

Group H

The theory assumed that the strength of a hypothesis increases when this hypothesis is confirmed, i.e., when given the stimulus presented, it is compatible with the outcome of the trial. If, for example, the stimulus presented on a given trial is large, horizontal . . . etc., and if the correct response to that stimulus is A, then the hypothesis according to which all large lines are As and all small ones are Bs is said to be confirmed on that trial, whether it is effectively being used by the S or not. Conversely, the strength of a hypothesis decreases when it is disconfirmed. It is therefore expected that the rating of a hypothesis on Trial *n* be an increasing function of the number of consecutive confirmations it has received on the immediately preceding trials; similarly, the rating of a hypothesis should be a decreasing function of the number of immediately preceding disconfirmations. The data shown in Fig. 1 support this prediction.

As an immediate consequence of this, the rating of the correct hypothesis, which is always confirmed, should, on the average, be higher than the rating of irrelevant hypotheses which, in turn, should be higher than the rating of the incorrect (always disconfirmed) hypothesis. These three classes of hypotheses are compared in Fig. 2 on the 10 trials preceding the criterion. The relation is as predicted on each of these trials. It can furthermore be observed that the rating of

Fig. 1. Group H, data prior to criterion: (a) Mean rating of a hypothesis as a function of the number *M* of preceding confirmations. (b) Mean rating of a hypothesis as a function of the number *M* of preceding disconfirmations.

Fig. 2. Group H: Mean rating of the correct, incorrect, and irrelevant hypotheses on each of the 10 trials immediately preceding the criterion.

the correct hypothesis increases over trials while the other ones decrease towards a common value.

The last question is concerned with the recency effect of positive and negative reinforcements. Table 1 shows the mean rating of a hypothesis on Trial n given some sequences of past reinforcements (where $+_n$ means that the hypothesis has been confirmed on Trial n , just before the moment of rating, $-_n$ means that it has been disconfirmed on Trial n , and similar notations are used for $n-1$ and $n-2$). The fact that Sequence a is better than b and that Sequence c is better than d illustrates a positive recency effect of reinforcements. Prediction of Group R from Group H

The theory furthermore assumes that the S is making one hypothesis at a time and is responding according to it. In this analysis, the probability of a hypothesis on a given trial will be supposed to be equal to the ratio of its strength to the sum of the strengths of all hypotheses. Using the average ratings of Group H at each trial and for each hypothesis as estimates of the corresponding strengths, thus enables one to compute the probability of each hypothesis on each trial. These probabilities are sufficient, knowing the stimulus presented, to compute the theoretical proportions of responses of either type in Group R on each trial. Fig. 3 shows the contingency diagram between the proportion of responses A inferred, as shown above, from the data of Group H, and the observed proportion, on each of

Table 1
Mean Rating of a Hypothesis on Trial n given the Sequence of its Reinforcements Respectively on Trials n and $n-1$, and $n-1$ and $n-2$

Sequence	Mean Rating	Sequence	Mean Rating
(a) $+_n, -_{n-1}$	2.46	(c) $+_{n-1}, -_{n-2}$	1.75
(b) $-_n, +_{n-1}$	0.77	(d) $-_{n-1}, +_{n-2}$	1.43

the 16 first trials of the experiment. The prediction is obviously relevant. A systematic bias must nevertheless be noticed in the sense of the low (high) values being underestimated (overestimated). The Spearman rank coefficient between observed and predicted values is .70.

CONCLUSIONS

The general assumption that the S was using hypotheses while responding in a concept identification task, and that learning basically affected the strength of these hypotheses was supported by two types of results: the sensitivity of ratings to positive and negative reinforcements, and the fairly good prediction of the data of Group R from the data of Group H.

Fig. 3. Observed proportion of responses A on each of the 16 first trials in Group R, and proportion predicted from the data of Group H.

REFERENCES

- BOWER, G. H., & TRABASSO, T. R. Concept identification. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Stanford: Stanford University Press, 1963. Pp. 32-94.
- FALMAGNE, R. Construction of a hypothesis model for concept identification. Preprint, 1968, University of Wisconsin.
- LEVINE, M. Mediating processes in humans at the outset of discrimination learning. *Psychological Review*, 1963, 70, 254-276.
- LEVINE, M. Hypothesis behavior by humans during discrimination learning. *Journal of Experimental Psychology*, 1966, 71, 331-338.
- RESTLE, F. The selection of strategies in cue learning. *Psychological Review*, 1962, 69, 329-343.
- RICHTER, M. Memory, choice and stimulus sequence in human discrimination learning. Unpublished doctoral dissertation, Indiana University, 1965.
- SUPPES, P., & SCHLAG-REY, M. Observable changes of hypotheses under positive reinforcement. *Science*, 1965, 18, 661-662.
- TRABASSO, T., & BOWER, G. H. Presolution dimensional shifts in concept identification: A test of the sampling with replacement axiom in all-or-none models. *Journal of Mathematical Psychology*, 1966, 3, 163-173.

NOTES

1. This study was sponsored by the Fonds National de la Recherche Scientifique, Brussels, Belgium.
2. Now at the Mental Health Research Institute, Ann Arbor, Michigan.
3. All the data involved are prior to the criterion.