Hypothesis sampling in successive solutions of a concept-ID task

VESTA S. GETTYS AND CHARLES F. GETTYS
OHIO STATE UNIVERSITY

The study examined changes in hypothesis sampling behavior as a function of eight successive solutions in a concept-ID task. It was found that the probability of a win-stay response increased with experience. No change· was observed in lose-stay behavior. The results are consistent with a sampling-with-partial-replacement hypothesis.

Several models have been proposed to describe and predict the sequence of responses which occur in a concept-ID task. All of the models assume that S's primary task is to learn which cue is relevant for the classification of the "correct" response. However, they differ in their assumptions concerning the way in which S comes to identify the relevant cue. Most of the models agree that if S makes a correct choice he will retain that cue on the next trial. However, they disagree about the sequence of events following an error.

Levine (1966) assumes that on the first trial S selects a hypothesis (H) at random from a set of Hs and tests to see if it is correct. If an error occurs, S eliminates the incorrect H and samples without replacement from the remaining set of available Hs. Bower & Trabasso (1963, 1964) assume that if an error occurs, S resamples with replacement from a pool of cues which *may* contain the H just sampled.

It is difficult to compare the results obtained by Levine with those of Bower and Trabasso because of differences in their experimental procedures. Two important differences seem to be the presence of non-reinforced trials in the Levine procedure and the differences in training procedures. Bower and Trabasso typically gave their Ss a brief, simple problem prior to the experimental session, and then ran them on a single experimental problem. Levine gave his Ss four preliminary problems which introduced them to the stimulus dimensions used in the experimental task, and also led them to expect trials in which they received no reinforcement. Thus, Bower and Trabasso obtained their data from naive Ss, while Levine's Ss were well trained. Also, Bower and Trabasso's Ss solved complete problems, while Levine's Ss only worked on a few four-trial sets.

It may be that the different conclusions reached by Bower & Trabasso (1963, 1966) and Levine (1966) are due to the relative naivete of their Ss. Perhaps S samples with replacement on his initial experience with the task, but on subsequent trials samples with partial or no replacement. To test these speculations, a task was designed which would reflect which H was being tested on any given trial, but which would

not require extensive pretraining. Results of the study may help to determine if some of the differences between results obtained by Bower and Trabasso and by Levine can be explained by changes in S's sampling behavior as a function of the amount of experience that he has had with the task.

Method

Sixteen psychology undergraduates served as Ss to fulfill a course requirement.

The stimuli consisted of four single digits typed in a diamond pattern on a 3 x 5 in. card. S's task was to determine whether a card was an "A" or "B". Each of the four positions on the card was defined as a dimension, and for a given solution one position was chosen to be the "relevant" dimension. Ss were told what values each position could take on; e.g., either a 1 or a 2 might appear in Position 1, a 3 or 4 in Position 2, etc. They were told to choose a position, and to multiply the number that appeared there by nine. (Multiplication was not necessary to the task as the odd- or evenness of the product was the same as that of the multiplicand. This procedure was adopted in an effort to conceal from S the purpose of the numerical part of the response.) In order to be consistently correct, they had to locate the relevant position, and then determine whether odd products or even products were As. (For example, if the stimulus card contained $7 \begin{smallmatrix} 1 \\ 3 \\ 6 \end{smallmatrix}$., and S was testing Position 2, his response might be "27-A.") The correct response, A or B, was printed on the back of the stimulus card, and, after responding, S turned the card over to find the correct response. The task was self-paced, and Ss continued to respond until they reached a criterion of eight trials in which the correct H was chosen. Each S solved eight such problems, one using each of the eight possible solution Hs. The order of solution of the problems was drawn from two 8 by 8 random Latin Squares. A record was kept of the stimulus card, response, and reinforcement for each trial.

Results and Discussion

In every two-trial sequence, there were four possible outcomes. If told "correct," S could retest the same H (win-stay) or shift to a new H (win-shift). If told "wrong," he could still retest the same H (lose-stay) or shift to a new H (lose-shift). The frequencies of win-stay, win-shift, lose-stay, and lose-shift responses prior to the last error were tabulated. The probability of occurrence of these events was calculated and is presented in Fig. 1.
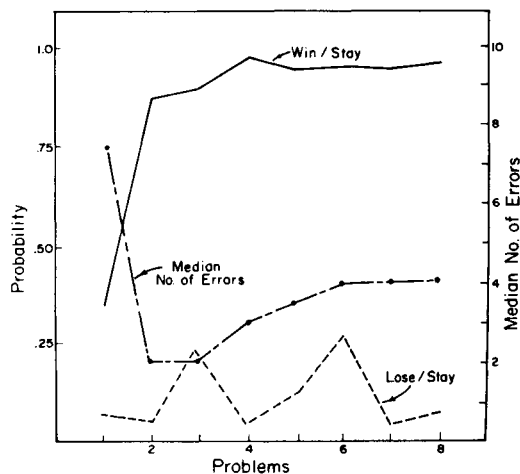
The probability of a win-stay changed as a function

Fig. 1. Changes in H testing and errors as a function of problems solved.

of practice. Both Bower & Trabasso (1963, 1966) and Levine (1966) predict a probability of 1.0 for a win-stay. In this study, it was only .35 in Problem 1 and increased to .93 for Problems 2-8. These data provide strong evidence that Ss change their H sampling behavior as a function of experience with the task ($\chi^2 = 199.1$, df = 1, p < 001).

Following an error, Levine (1966) found that his Ss retained the same H with a probability of .05. Bower & Trabasso (1963) predict that if Ss are sampling at random from eight possible Hs, the probability of retesting the same H on the next trial would be .125. In this study the probability of a lose-stay was .067 ±.058 (a 95% confidence interval) for Solution 1 and .102±.03 for Solutions 2-8. For Solution 1, the 95% confidence limits include both the .05 probability obtained by Levine and the .125 probability predicted by Bower and Trabasso. This measure does not differentiate between the two models. Levine's results are outside of the 95% confidence limits obtained for Solutions 2-8, and the obtained probability is close to sampling-with-replacement predictions. This indicates that Ss are sampling with at least partial replacement after experience with the task, as Trabasso & Bower (1966) suggest.

The median number of errors for each of the eight problems is also presented in Fig. 1. The median was 7.5 errors for Solution 1 and about 3.5 for Solutions 2-8. The mean trial of the last error for Solution 1 was 13.56, but for Solutions 2-8 it decreased to about 7.0. This indicates that most of the learning "set" took place in the first solution.

The protocols of the 16 Ss were checked for types of errors. On Solution 1, eight of the protocols had one or more retests of a previously tested H, and the average for Solutions 2-8 was 4.4. Over all, a total of 39 of the 128 response protocols (30%) contained instances of S's retesting a previously tested H. In 19 cases, Ss retested the same H at least twice after being told "wrong." There were also eight cases of Ss who tested the correct H, but failed to condition.

The principal point to be made in this study is that differences between the results of Bower & Trabasso (1963, 1966) and Levine (1966) may be due to differences in experience of Ss with concept-ID tasks. Both Trabasso & Bower (1966) and Levine seem to be wrong in their assumption that the win-stay probability for all Ss is 1.0. Results of this study indicate that while the probability of win-stay behavior increases as a function of practice, it does not reach 1.0. After experience with the task, the probability of S's resampling a previously tested H is closer to the Trabasso & Bower (1966) sampling-with-partial-replacement predictions.

Results of this study indicate that Ss do change their H sampling behavior as a function of experience with a concept-ID task, and that studying problem-solving behavior after varying amounts of practice can lead to different conclusions.

References

BOWER, G., & TRABASSO, T. Reversals prior to solution in concept identification. *J. exp. Psychol.*, 1963, 66, 409-418.
LEVINE, M. Hypothesis behavior by humans during discrimination learning. *J. exp. Psychol.*, 1966, 71, 331-338.
TRABASSO, T., & BOWER, G. Presolution reversal and dimensional shifts in concept identification. *J. exp. Psychol.*, 1964, 67, 398-399.
TRABASSO, T., & BOWER, G. Presolution dimensional shifts in concept identification: A test of the sampling with replacement axiom in all-or-none models. *J. Math. Psychol.*, 1966, 3, 163-173.