The response shift phenomenon in discrimination learning set [1]

ROBERT E. BOWMAN AND KEN-ICHI TAKEMURA

THE UNIVERSITY OF WISCONSIN REGIONAL PRIMATE RESEARCH CENTER

A response-shift effect was exhibited by monkeys trained on a modified discrimination learning set, under conditions in which neither object to be discriminated had been previously displaced. This example of the effect therefore could not depend on any tendency of the monkeys to investigate non-displaced objects. It is suggested that the present effect might have its origin in a postulated tendency of monkeys to retain irrelevant but rewarded first trial Hypotheses compared to a tendency to replace irrelevant, nonrewarded first trial Hypotheses.

Moss & Harlow (1947) first noted the tendency of monkeys to make more errors following a Trial 1 reward than following a Trial 1 nonreward in object discrimination learning set (ODLS). This tendency was later termed "response shift" and was given the status of an "error factor" by Harlow (i.e., an error-producing response pattern). It was also theorized (e.g., Harlow, 1959) to result from a postulated tendency of the monkey to explore or investigate both objects in a discrimination problem.

The purpose of this paper is to present an instance of the response shift phenomenon obtained under conditions in which the exploratory tendency could not operate. This instance occurred unexpectedly in an experiment dealing with hypothesis learning in the monkey (to be reported elsewhere), and has led to the formulation of an alternate explanation of the response shift effect.

Method

Eight rhesus monkeys, 2 to 4 years of age, and without prior ODLS, were tested in the Wisconsin General Test Apparatus or WGTA (Harlow, 1949), using a formboard with two foodwells 12 in. apart. Ss were extensively adapted to take food from the formboard and were then pretrained to displace objects using insoluble problems.

Next, Ss were trained for 32 sessions, each of eight six-trial discrimination problems (Phase I). The problems were designed to train one group of 4 Ss (Group R) on the Hypothesis or H (Levine, 1959, 1963) of "win-stay-on-object" and another group of 4 Ss (Group N) on "lose-shift-to-other-object." To do this, two new objects were introduced on Trial 1 of each problem. One of these objects then recurred on each of the following five trials, paired each time with a new object. For Group R the recurring object of each problem was always rewarded and the nonrecurring objects were not rewarded. For Group N, the recurring object of each problem was never rewarded, and the nonrecurring objects were always rewarded. Within each group, two further treatment conditions were differentiated, namely
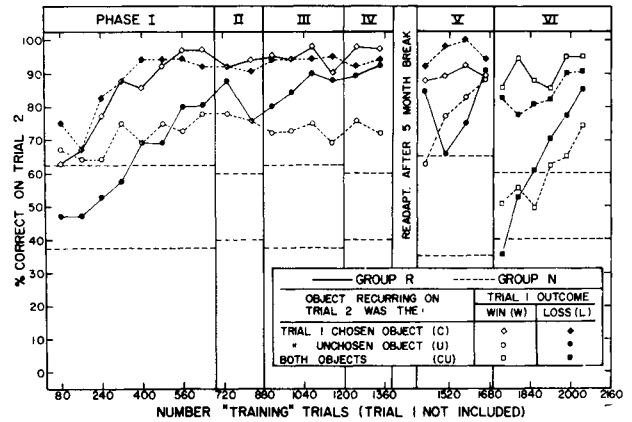

Fig. 1. Performance of Trial 2. throughout the experiment.

whether on Trial 1 the S was rewarded (i.e., "won," W) or not (i.e., "lost," L), and whether the recurring object was that chosen (C) or unchosen (U) on Trial 1. By this training procedure, the correct Hs of "win-stay" (Group R) or "lose-shift" (Group N) could be utilized on Trial 2, given a C condition. Additionally, correct Hs of "approach-recurring-object" (Group R) or "avoid-recurring-object" (Group N) could be used given either a C or a U condition.

Subsequently, the Ss were trained for 16 sessions (Phase II) exactly as in Phase I except that test trials were substituted for training trials on six of the Trial 2's and six of the Trial 6's within each 48-Trial session. On these test trials, both objects were baited and the reverse of the usual object recurred paired with a new object. A new problem was begun immediately following each test trial. Ss were next retrained for 16 sessions (Phase III) exactly as in Phase I. Phase IV, of 16 sessions, was then presented exactly as was Phase II except that on the test trials both objects of the immediately preceding trial recurred. The results on these Phase II and IV test trials will be reported elsewhere. Finally, following a break of 5 months, the Ss were retrained for 16 sessions in Phase V as in Phase I. Then 24 sessions were presented (Phase VI), each consisting entirely of 16 three-trial problems of the usual ODLS type (both objects of Trial 1 of a problem recurring on both subsequent trials).

Results

The relevant results consist of the Trial 2 performance on the training trials (Fig. 1). The two upper curves (rhombic points) for Phases I-V represent C conditions, in which two correct Hs were simultaneously

available to each group. Performance quickly approached maximum, and response shift appeared only at the very beginning of the C conditions, in the form of a slight performance superiority following a nonrewarded Trial 1.

The two lower curves (circle points) of Phases I-V represent the U conditions, in which only one correct H was available to each group. Those animals rewarded on Trial 1 (win) performed at 60-70% correct throughout, whereas those animals not rewarded on Trial 1 (loss) improved from chance to 90% correct by the end of Phase IV. Response shift appears as the difference between the solid circle and open circle curves from Phases II through IV. Despite the fact that these are between-group curves, these data parallel those of previous studies using regular ODLS (e.g., Davis et al, 1953; Riopelle, 1953; Behar, 1961).

It is clear that the present response shift cannot be explained as an exploratory tendency toward the non-displaced object, since in the U conditions *neither* of the two objects presented to the monkey on Trial 2 had been previously displaced. Accordingly, a new explanation for response shift has been devised which is general enough to be applicable both to the present effect and to response shift observed in regular ODLS.

First, assume that the monkey enters Trial 1 of a problem with some H, where the H is defined as a mediating process determining the response. Next, assume that if reward is obtained, the H is confirmed and therefore retained, but if reward is not obtained, the H is rejected and another H chosen. Such processes have been reported (Levine, 1964) for the human.

Now, consider the case in which the Trial 1 H is irrelevant. If it is rewarded and therefore retained, the monkey will make 50% errors on Trial 2. However, if it is not rewarded and is therefore rejected, the monkey will choose a new H. As learning progresses, this new H will tend to be the correct one. Thus, the monkey will tend to make fewer errors on Trial 2. These processes may be summarized by saying that the monkey retains error factors given a Trial 1 reward, but learns to abandon error factors for the correct H given a Trial 1 nonreward.

If the above explanation is valid, then only those instances in which the monkeys begin a problem with an irrelevant H (an error factor) would give rise to the response shift effect. Furthermore, this effect cannot begin to appear until the correct H or Hs begin to be available for choice on Trial 2, accounting for the absence of response shift early in ODLS. Additionally, response shift could only be manifested so long as the monkeys continued to use irrelevant Hs on Trial 1. This continued use of irrelevant Hs seems possible since the correct Hs depend on the reward outcome (or

also in the present experiment on the object recurrance) and the correct Hs are therefore not available prior to the Trial 1 choice.

From this line of reasoning, the development of learning set could be regarded as requiring the suppression of Trial 1, preoutcome Hs, and the response shift data then suggest that this suppression occurs most readily if nonreward occurs on Trial 1. If such suppression finally began to occur following Trial 1 reward, then performance following rewarded first trials would approach maximum and the response shift phenomenon would disappear.

This last event occurred in the present data (Fig. 1, Phase V, dotted line-open circle curve) when performance following a *rewarded* Trial 1 increased to a final level in excess of 90% correct. (The drop and recovery of the concurrent solid line-solid circle curve appeared to represent the adverse reaction of two animals to a change in experimenters.)

In the following Phase VI, in which both objects of a problem recurred on every trial, the Group R monkeys now learned "lose-shift-to-other-object" at about the same rate as the Group N monkeys learned "win-stay-on-object," and although the response shift tendency reappeared, it was not statistically significant. Furthermore, the upper pair of curves in Phase VI, for the conditions permitting the monkeys to utilize their previously learned Hs ("win-stay-on-object" for Group R and "lose-shift-to-other-object" for Group N) exhibited no significant difference and no response shift effect. These monkeys therefore eventually learned to deal with Trial 1 reward, and the response shift effect appeared to be a transitory phenomenon in the development of the learning set.

References

Behar, I. Learned avoidance of nonreward. *Psychol. Rep.*, 1961, 9, 43-52.

Davis, R. T., McDowell, A. A., & Thorson, N. Four-trial object-quality discrimination by monkeys. *Proc. S. D. Acad. Sci.*, 1953, 32, 132-142.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 2., New York: McGraw-Hill, 1959.

Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, 56, 51-65.

Levine, M. A model of hypothesis behavior in discrimination learnset. *Psychol. Rev.*, 1959, 66, 353-366.

Levine, M. Mediating processes in humans at the outset of discrimination learning. *Psychol. Rev.*, 1963, 70, 254-276.

Levine, M. Hypothesis behavior by humans during discrimination learning. Paper read at Psychonomic Society Convention, 1964.

Moss, Eileen M., & Harlow, H. F. The role of reward in discrimination learning in monkeys. *J. comp. physiol. Psychol.*, 1947, 49, 333-342.

Riopelle, A. J. Transfer suppression and learning sets. *J. comp. physiol. Psychol.*, 1953, 46, 108-114.