Psychological Reports, 1963, 12, 359-362.

SECONDARY REINFORCEMENT AND FRUSTRATION

GORDON H. BOWER[1]
Stanford University

Summary.—It was pointed out that conventional tests for secondary reinforcement are also theoretically optimal conditions for the production of aversive frustration. Hence, positive demonstrations of secondary reinforcement create a paradox for frustration theory. A resolution of this paradox is offered in terms of the chain of behavior cued off by the secondary reinforcing stimulus. The analysis implies that secondary reward effects will be maintained longer the longer is the chain of behavior cued off by stimuli presented following the to-be-learned response. To test this prediction, rats were first trained to run an alley for food reward. Subsequently, they learned to press a lever to get out of the start box; *Ss* in the Short-chain condition exited directly into the goal box, whereas *Ss* in the Long-chain condition ran through the alley to the goal box. Sixty tests were given with food reward omitted from the goal box. As predicted, the bar pressing performance of the Long-chain *Ss* was consistently superior to that of the Short-chain *Ss*.

Within the context of modern Hullian theory, there is an apparent conflict between the notions of frustration and secondary reinforcement. According to Amsel's hypotheses (Amsel, 1958), nonreward occurring when *S* is expecting reward is assumed to produce a frustration reaction. The term "expecting reward" is identified in the theory with the occurrence of a fractional anticipatory goal response, r_g, that has become conditioned to cues preceding the primary consummatory reaction. Frustration is considered to be aversive, having effects much like those ascribed to punishment. For example, it is the anticipation of aversive frustration that Amsel supposes is one of the factors producing extinction of previously rewarded (now unrewarded) instrumental responses.

Consider now a conventional paradigm for a study of secondary reinforcement. During the initial phase of the experiment, a buzzer is paired repeatedly with the delivery of food which the hungry animal eats. A test for secondary reinforcement is then provided by inserting a lever into the cage; when the lever is depressed, the buzzer is briefly sounded but no food is delivered. The conventional finding is that *S* presses the lever to sound the buzzer more frequently than will a control *S* not given the initial pairings between buzzer and food reward. An attractive hypothesis offered by Hull (1943) identifies secondary reinforcement with the elicitation of r_g which has been conditioned to the buzzer. This hypothesis is useful in that it suggests that the phenomena of secondary reinforcement follow the dynamic laws of classical conditioning. To the extent that this program has been tested, it seems to hold up fairly well (Kimble, 1961).

At one level of analysis, however, this demonstration of secondary rein-

[1]Ray Sweigert assisted in carrying out this experiment. Support for this research was provided by a Grant M-3849 to the author from the National Institutes of Mental Health.

forcement raises difficulties for Amsel's frustration theory. The difficulty is that the method of testing for secondary reinforcement should also be an optimal condition for producing frustration; the similarity of the two operations was recognized previously by Amsel (1961). The lever press sounds the buzzer, which evokes r_g; but since no reward is forthcoming on the test, frustration should result. The question is, why should an animal want to put itself into a stimulus situation that is consistently frustrating and aversive? There are several ways in the theory to resolve this "paradox"; this paper shows one line of argument and reports an experiment which tends to support it.

First, it is oversimplifying matters to suppose that the buzzer CS is conditioned only to r_g. In most cases, the CS is a discriminative stimulus which sets off a chain of instrumental behaviors eventuating in the consummatory response. For example, the CS may elicit orientation toward the food cup, approaching it, placing paws on it, thrusting the head into the food cup, and seizing the food pellet. Schoenfeld, Antonitis, and Bersh (1950) have emphasized this sort of chained behavior that is cued off by the secondary reinforcing stimulus. Presumably, the buzzer CS and response-produced cues from early members of this stable chain become connected to r_g. According to Hull's hypothesis, it is the elicitation of this r_g which reinforces new behavior (lever pressing) that produces the buzzer. Secondly, on nonrewarded trials the frustration reaction is evoked at the end of this chain of behavior, when *S* thrusts its head into the food cup and finds no food. By Amsel's theory, part of this frustration reaction will become conditioned to those cues preceding its elicitation. A number of trials will be required for this conditioning to take place, and further time is required for the conditioned frustration reaction to move back along the behavior chain evoked by the CS. We will assume that the buzzer will continue to exert a reinforcing effect upon prior lever presses until that point in the test series at which the early members of the response chain (to the CS) are disrupted due to the anticipation of frustration.

This analysis admittedly is complicated, but several implications follow from it. This paper reports a test of one of these implications, that relating to the effect of varying the length of the chain of behavior cued off by the CS. The longer is this chain of behavior, the more delayed is the aversive frustration reaction following the test response which produces the nonrewarded CS. Furthermore, with a longer chain, more unreinforced trials will be required for the effects of frustration to work back and disrupt the beginning members of the chain. Both of these factors lead to the prediction that the learning and maintenance of performance of a new test response will be better the longer is the chain of behavior between the presumptive secondary reinforcing stimulus and the presumptive point of frustration due to nonreward.

METHOD

Subjects.—Fifteen male albino rats, 90 days old, were housed in individual living cages. They were maintained on 10 gm. daily of Purina lab checkers (given following their daily experimental runs) and free access to water in their home cages.

Apparatus.—A black start box, 12 × 6 × 6 in. deep, adjoined a gray runway. The runway involved four right-angle turns in the order left, right, right, left and was 48 in. long and 6 × 6 in. high. This runway led into a white goal box 12 in. long and 6 × 6 in. high, which had a food cup on the floor at the rear of the box. All parts were constructed of wood and covered with hardware cloth. The start box and goal box had guillotine doors which were used to confine *S*.

In the second part of the experiment, a lever was inserted through one wall of the start box, near the exit door. The lever paddle was 2 × 3 in. and a depression of .25 in. was sufficient to make contact. A flashlight bulb outside the runway indicated to *E* when the lever was depressed. During testing, the start door was raised when *S* pressed the lever.

Procedure.—There were 86 initial trials of training *S*s to run through the runway to the goal box for a reward of a small .2-gm. pellet of wet mash. During the first 56 of these trials, all trials were rewarded. During the last 30 trials, 15 trials were rewarded and 15 were nonrewarded in random order. There were six trials per day with at least a 4-min. interval between trials for a given *S*. During these trials *S* was timed with a manual stop watch (to .1 sec.) from the opening of the start door to his full-body entrance into the goal box. After this initial training *S*s were divided into two groups of 8 and 7 for the Long and Short test conditions. For the test phase, the lever was inserted into the start box and *S* was required to press the lever to open the start door. In the Long condition the start door opened into the gray runway which *S* could run to the white goal box. In the Short condition the runway was removed and the goal box was adjoined to the start box; depression of the lever opened the start door which exited directly into the white goal box. In both conditions, food reward was no longer present in the goal box during these tests. On test trials *S* was placed at the rear end of the start box facing the door. A stop watch was started manually when *S* was released. When *S* depressed the lever, the start door was raised and simultaneously one hand of the stop watch was stopped and a second hand started. When *S* entered the goal box, the second hand of the watch was stopped. Thus, measures were obtained of latency of the bar pressing response and time to go from the start box to the goal box. These measures were converted to reciprocals (speeds). Each trial terminated with *S* in the nonreward goal box for 10 sec. If *S* had not reached the goal box within 60 sec. after the trial started, he was picked up and placed in the goal box. Each *S* received 60 test trials, 6 a day for 10 days. Two *S*s in the Short condition died during testing, leaving 8 *S*s in the Long and 5 *S*s in the Short condition.

Results and Discussion

The results of primary interest concern the learning of the new bar pressing habit by *S*s in the Long and Short test conditions. The prediction is that *S*s in the Long condition will learn the habit better since the effective point of frustrative nonreward for them is farther removed from the occurrence of the bar press response. The major results on the average speed of bar pressing in the two conditions are shown in Fig. 1, plotted in 10 daily blocks of 6 test trials. Although the two curves begin relatively close together, they diverge over test days with the Long condition becoming clearly superior. The means are not significantly different from one another during the first three test days, but thereafter the differences are large and statistically significant. Over the last 7 days of testing, the mean speeds were 247 and 121 for the Long and Short conditions, respectively. These are reliably different ($t = 2.51$, $df = 13$, $P < .05$). Thus, the prediction is confirmed.

Concerning speeds from the start box to the goal box during testing, no useful comparisons can be made between groups because of the different lengths of chain. Both groups showed a slight increase in average speeds over early

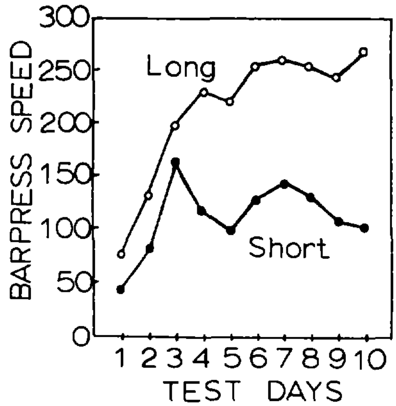

FIG. 1. Group average speeds of the bar press response during the 10 days of testing

test trials, apparently adapting to the novelty of the new procedure (bar pressing) for opening the start door. Thereafter, speeds from start to goal box began to decline under the influence of nonreward. However, all *Ss* were still running to the goal box within 60 sec. at the end of the 60 test trials.

The procedure was clearly satisfactory for producing learning and persistence of the new bar pressing habit. Our purpose in giving partial reinforcement at the end of runway training was to build in some persistence to the response of approaching the goal box so that the Long-Short variable could be seen operating over a number of test trials. The design succeeded almost too well in this respect since most *Ss* in the Long condition persisted with rapid bar-press responses throughout the 60 test trials. Presumably, with more extensive series of nonrewarded trials, the bar pressing would extinguish as the runway performance declined farther. However, in this situation there probably is some low residual reinforcement for bar pressing simply because it is followed by escape from the start box (McNamara & Paige, 1962).

The results of this experiment were predicted from a particular theoretical analysis, namely, that with long response chains it is more difficult for frustrative effects of nonreward to act back upon the new response which produces the secondary reinforcing stimuli. This hypothesis would also predict that starting speeds would be more resistant to extinction the longer is the runway that *S* traverses during training and extinction. It is possible, of course, to explain the present results in other terms, e.g., the cue sequence in the Long condition produces greater cumulative secondary reinforcement. However, our purpose here was not to rule out alternative *ad hoc* explanations, but rather to test a clear implication of one resolution of the secondary reinforcement-frustration paradox. For this limited aim, the results are satisfactory.

REFERENCES

AMSEL, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.*, 1958, 53, 102-119.

AMSEL, A. Review of O. H. Mowrer's *Learning theory and behavior. Contemporary Psychol.*, 1961, 6, 33-36.

HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

KIMBLE, G. A. *Hilgard and Marquis' Conditioning and Learning.* New York: Appleton-Century-Crofts, 1961.

MCNAMARA, H. J., & PAIGE, A. B. An elaboration of Zimmerman's procedure for demonstrating durable secondary reinforcement. *Psychol. Rep.*, 1962, 11, 801-803.

SCHOENFELD, W. N., ANTONITIS, J. J., & BERSH, P. J. A preliminary study of training conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1950, 40, 40-45.

Accepted February 11, 1963.