
LEARNING IN THE RAT (*RATTUS NORVEGICUS*) UNDER POSITIVE VS. NEGATIVE REINFORCEMENT WITH INCENTIVE CONDITIONS CONTROLLED

REED LAWSON AND LUKE S. WATSON, JR.¹

Department of Psychology, The Ohio State University, Columbus 10

A popular stereotype holds that psychologists are opposed to the use of punishment as a device for teaching organisms. As with all stereotypes, this is an oversimplification, but there is some truth to it. Some very definitive experiments (Estes, 1944) have shown that punishment by itself does not permanently weaken the tendency to perform a response. But there are other ways of using noxious stimulation in teaching, and neither psychologists nor the public are entirely convinced that these may not be efficient methods of behavioral control (leaving out ethical considerations, of course).

One way of using punishment is to administer it if a response is *not* performed. A warning stimulus is presented and then, if the appropriate response is not performed, a noxious stimulus is given. In the psychologist's vernacular this is called "avoidance learning," i.e., the subject learns to avoid punishment by performing a particular response under specified stimulus conditions. Technically, this is called learning with *negative reinforcement*: a response is strengthened (reinforced) by the removal or the prevention of a noxious stimulus. Both in the laboratory and in everyday life people are frequently impressed with the apparent ease with which behavior can be controlled by such a technique; avoidance learning is a device much used in the laboratory when psychologists ostensibly seek to establish a habit quickly.

The alternative to teaching via negative reinforcement is to use reward, or *positive reinforcement*. Instead of removing, or preventing the occurrence of some

¹Based on a thesis submitted by the second author to the Graduate School of The Ohio State University in partial fulfillment of the requirements for the M.A. degree.

noxious stimulus when the organism performs acceptably, an additional, positive event is made to occur when the subject makes the desired response. Everyone knows that positive reinforcement works, too. The unresolved question is whether it works as efficiently as negative reinforcement.

Aside from a few poorly controlled studies in formal educational settings, there have been no direct comparisons of the efficacy of positive and negative reinforcement as teaching methods. One reason for this is that not all learning situations have been conducted in directly comparable fashions with only the type of reinforcement varying. For instance, it has been asserted that the learning situation used in the present experiment is difficult for a rat to learn under negative reinforcement conditions (Meyer et al., 1960). But an exactly comparable presentation of stimulus and reinforcement time relations, with a comparable response latency requirement, is not to be found in the research literature. Perhaps this situation produces equally poor learning when positive reinforcement is used.

Another difficulty is that of equating the incentive value of positive and negative reinforcers. Suppose that exactly the same learning situation was set up for two groups of organisms, one group getting positive and the other group getting negative reinforcement. No matter which group did better than the other, how could we be sure that the incentives for learning were approximately equivalent for the two groups?

The present experiment is an attempt to handle both of these methodological difficulties and make a direct comparison of the effectiveness of positive and negative reinforcers in learning. The general hypothesis tested in this experiment was that positive reinforcement is a more efficient method of developing a habit than is negative reinforcement, even when the incentive value of the positive reinforcer is known to be extremely low. This hypothesis stems from the general proposition that aversive stimulation is not as effective as positive reinforcement in controlling behavior (e.g., Skinner, 1954).

The approach used to test this hypothesis was to use three groups of laboratory rats: one group receives a standard form of positive reinforcement, one group a standard degree of negative reinforcement, and a third group is given a decidedly inferior form of positive reinforcement. By "standard" reinforcements is meant those conditions that have repeatedly been found to be effective reinforcement conditions for the establishment of habits in the laboratory using the hooded rat. The learning situation used was one which has been extensively studied using negative reinforcement, but not exactly duplicated previously using positive reinforcement.

METHOD

The general procedure used in this experiment was as follows. Laboratory rats were placed in a small, insulated cubicle which contained a lever, a food trough, and a grid floor suitable for administering electric shocks to the feet. All rats were taught the same problem: to press the lever whenever an auditory signal sounded in the box. For two groups, every such response was followed by the presentation of a food pellet; for one of these groups the pellet was a standard, highly acceptable food, for the other it was a special pellet just barely acceptable to a hungry rat. For the third group, failure to press the lever within 5 sec after the tone came on resulted in an electric shock. A response during the 5 sec period prevented the occurrence of shock; if the shock came on, a lever-press was still necessary to terminate the shock. In all groups, the first response in the presence of the tone ended the tone and began the timing of the intertrial interval, as well as providing the appropriate reinforcement. Subjects were given a fixed number of trials per day, and the criterion was the percentage of responses made within 5 sec. of the onset of the tone. The procedural details were as follows.

Subjects and Apparatus

Thirty experimentally naive male and female hooded rats, between the ages of 90 and 240 days with a mean age of 151 days and a median age of 125 days were selected from the colony housed at The Ohio State University Laboratory of Comparative and Physiological Psychology. Their weights ranged from 170 to 418 g, with a mean weight of 277 g and a median of 267 g. All groups were equally diverse as to age, weight, and sex.

The apparatus consisted of a commercially-produced (Foringer & Co.) free operant conditioning box with a test compartment $9\frac{7}{8}$ in. wide, $10\frac{5}{8}$ in. long, $11\frac{3}{4}$ in. deep. A stainless steel closed tubular bar which was 2 in. long and $\frac{1}{2}$ in. in diameter protruded $\frac{3}{4}$ in. from the front wall of the compartment, its center located $3\frac{1}{4}$ in. above a grid floor and 2 in. from the wall of the box. The pellet trough was $1\frac{3}{4}$ in. above the floor and 4 in. to the left of bar center. All

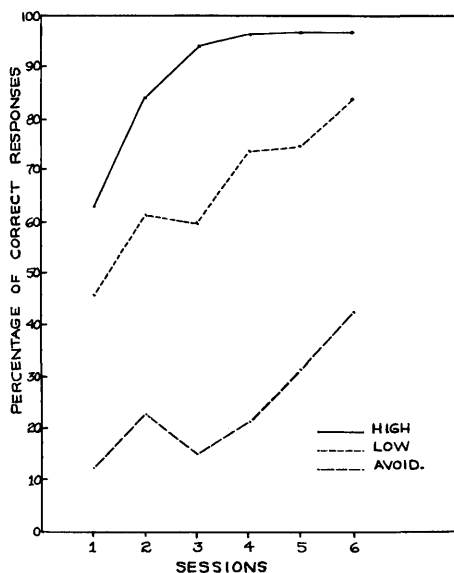


FIGURE 1. Mean percentage of correct responses per discrimination training session for each group.

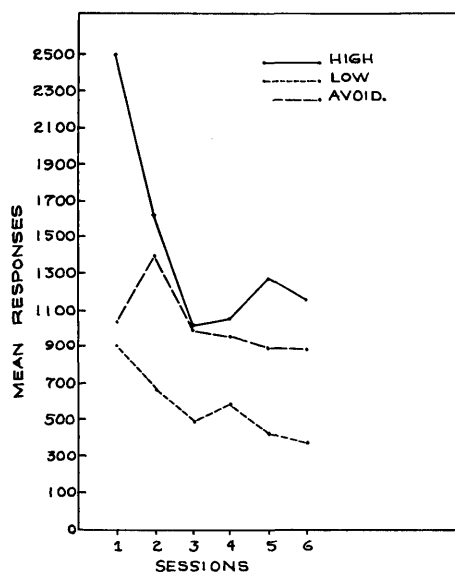


FIGURE 2. Mean number of total responses per discrimination training session for each group.

stimulus and reinforcing events in the box were programmed with commercially-produced equipment designed for that purpose. Responses were recorded simultaneously on a Gerbrands cumulative recorder and on impulse counters. The electric shock used in the negative reinforcement (Avoidance) group was 270 v.d.c. at 1.1 ma. The auditory signal used was a complex tone with a fundamental of 800 cycle/sec.

The high incentive reward (which will be called simply High Reward) was a commercially-produced (P. J. Noyes Co.) lab rat food tablet, 4 mm x 3.3 mm x 45 mg. It consisted of animal feed, bleached flour, dry milk solids, glucose solids, gelatin, and calcium phosphate. Rats are known to accept this pellet readily, and will even eat it in preference to their standard diet. The low incentive reward (Low Reward) was a pellet of the same size made up of three parts sucrose, one part lactose and small amount of lemon flavored Koolade mix. In a free-choice eating situation, rats ate 6.5 times as many High Reward pellets as Low Reward pellets. In a free-responding situation, rats responded at a rate 2.8 times as fast

when receiving High Reward as when receiving Low Reward. There is no doubt that the pellets differ in incentive value, and qualitative evidence suggested that the Low Reward pellet was close to the bottom of a hypothetical scale of acceptability of rewards.

Procedure

As stated before, there were three groups in this experiment: an avoidance group receiving shock as a negative reinforcer, a High Reward group and a Low Reward group. Other than type of reinforcement all subjects were given experimental treatments as identical as possible. Subjects were run in replications of three. During one replication, the programming equipment failed and so this group of three subjects was discarded, leaving 27 subjects—9 per group. The three subjects in any replication were housed in a common cage with a constant water supply. For 10 days prior to training all subjects were put on a 23-hr food deprivation schedule, and fed for 60 min per day to reduce their weight to approximately 80 per cent of normal.

Preliminary training.—Each subject in the reward groups was first trained to approach the food trough as soon as a pellet was delivered by giving 60 automatic reinforcements on an irregular interval averaging 30 sec. This was followed the next day by shaping them to press the lever, which continued until each subject had made 60 successful (reinforced) responses itself. The subjects in the avoidance group were trained to turn off the shock by pressing the lever. Shocks were delivered every 5 to 8 sec and continued until these subjects had made 60 successful escape responses.

Discrimination training.—Having established the response, the main part of the experiment began—bringing the response under the control of an external stimulus. The contingency set up for all subjects was that they depress the lever within 5 sec after the onset of the tone in order for the response to be considered correct. Correct responses (i.e., those meeting this criterion) were automatically recorded. The Avoidance group received a shock 5 sec after the onset of the tone unless the lever was pressed, and shock continued until the subject made this response. The Reward groups received one pellet after they pressed the lever at any time following the onset of the tone (i.e., they were not “punished” by no food for failing to respond quickly enough). In all cases the first response after the tone came on terminated the tone. Responses when the tone was not sounding were never reinforced and they did not affect the program in any way. Tones were presented every 60 sec after the termination of the last trial. All subjects received six discrimination training sessions, each of which was 2 hr long. The values of the time between trials, duration of the external stimulus, and intensity of the shock were all determined by extensive preliminary work in our laboratory to produce the best avoidance conditioning for our strain of rats in our type of apparatus.

RESULTS AND DISCUSSION

Figure 1 shows the mean percentage of correct responses (responses made in less than 5 sec after the onset of the tone) for each group over the six discrimination training sessions. Analysis of variance revealed that the difference between overall means was significant ($F=64.85$, 2 & 24 df, $p<0.001$); individual t -tests taking the groups two at a time showed that all groups were significantly different from each other (all t 's significant beyond the 0.01 level). Furthermore, the interaction between type of reinforcement and sessions was significant ($F=2.12$, 10 & 120 df, $p<0.05$), indicating that the slopes of the learning curves for the three groups are also different.

As mentioned at the outset, there is the theoretical possibility that the negative reinforcer is not as great in incentive value as the positive reinforcers. One indi-

cation of motivation in a free responding situation is overall rate of responding (not just correct responses). Figure 2 shows the mean number of total responses emitted by each group during each session. The *F* for overall differences was again significant (4.27, 2 & 24 df, $p < 0.05$), but individual *t*-tests showed that this was due entirely to the fact that the Low Reward group was much less motivated to respond than the High Reward group. In terms of average number of responses, the Avoidance group did not differ from the High Reward group. Since all subjects are equally hungry, presumably the only difference in motivational factors would be due to the incentive value of the reinforcement available in the situation. By this reasoning, the inferiority of the learning of the Avoidance group cannot be attributed to lower motivation.

Within the context of at least one very common kind of laboratory learning situation, then, no evidence has been found to support the possibility that negative reinforcement (the removal of the threat of physical punishment) is a particularly effective way of obtaining learning. Learning does occur, but far more slowly than when positive reinforcement (reward) is used. Even an extremely weak reward is better than punishment.

It must be pointed out that there are laboratory learning situations in which discriminated avoidance learning occurs faster and reaches a higher level of accuracy than in the bar-pressing situation. But this is also true for responding that is positively reinforced. Direct comparisons of positive and negative reinforcement are lacking in these other learning situations as well, and until these are made the complete generality of the present findings is, of course, unknown.

Perhaps it should be pointed out in conclusion that the Reward groups were, in a sense, under far less impetus to respond within 5 sec of the onset of the tone than the Avoidance group; a food pellet would be given at any time after the tone sounded. Nevertheless, these subjects reached the criterion much sooner and more often than the subjects for whom the 5 sec criterion should have been more "obvious".

SUMMARY

Three groups of rats were trained according to the procedure for developing discriminated bar-press avoidance behavior, except that two groups received positive reinforcement instead of shock escape or avoidance for responding. One positive reinforcement group received a highly acceptable food reward while the other was reinforced with a pellet of minimal acceptability. Both positive reinforcement groups were superior to the avoidance group in rate of learning and final level of performance.

ACKNOWLEDGMENT

We wish to thank Dr. G. P. Guth of The Ohio State College of Pharmacy for his assistance in preparing and making the low incentive pellets used in this experiment.

REFERENCES

- Estes, W. K.** 1944. An experimental study of punishment. *Psychol. Monogr.* 57(263): 40 pp.
Meyer, D. R., C. Cho, and Ann F. Wesemann. 1960. On problems of conditioning discriminated lever-press avoidance responses. *Psychol. Rev.* 67: 224-228.
Skinner, B. F. 1954. *Science and human behavior.* MacMillan Co., New York pp. 1-461.
-