

Reward magnitude shifts: A savings effect¹

ALLAN R. WAGNER AND EARL THOMAS

YALE UNIVERSITY

Two groups of rats were first trained to traverse a runway, one with large and the other with small food rewards. Both groups were then given additional training with small rewards until their performance equalized. In final training, both groups were shifted to the larger rewards. The group with a prior history of large rewards, evidenced savings in the final stage of training, in terms of a faster rate of performance change to the level appropriate to the larger reward.

Under certain conditions the magnitude of reward variable can be shown to have relatively persistent effects upon the runway performance of the rat. Thus, Ss trained with large, as compared to small, rewards when subjected to experimental extinction will continue to evidence their different reward histories over the course of extinction (e.g., Wagner, 1961; Hulse, 1958). Under other conditions, however, reward magnitude appears to have a less persistent influence. Thus, Ss trained with large rewards and then shifted to smaller rewards will, after a relatively small number of postshift trials, come to respond at speeds indistinguishable from those of Ss trained only with small rewards (e.g., Crespi, 1944; Zeaman, 1949). A question which this poses is whether shifting to a new, but nonzero, value of reward is an especially effective method of eliminating any influences of prior reward magnitudes, or whether relatively persistent historical influences exist but are simply not observed in performance under these conditions? The results of the present investigation of shifts in reward magnitude support the latter interpretation.

Method

The Ss were 40 male albino rats, 90 to 110 days of age at the beginning of the experiment. They were randomly assigned to two groups of 20 Ss each.

The apparatus was a straight, enclosed alley 3 in wide and 4 in high throughout, divided by guillotine doors into a 12-in start box, a 45-in runway, and a 15-in goal box. A 1-in deep food cup was attached to the end wall of the goal box with its rim 1.5 in from the floor. Total-Response time was measured from the opening of the start box door until the interruption of a photocell beam just prior to the rim of the food cup, and was fractionated by means of intermediate photocells located 12 in and 48 in from the start box, into three successive segments, start-time, running-time, and goal box-time.

Ten days prior to experimental training, Ss were placed on a 24-hour food deprivation schedule and brought to between 75% and 80% of their ad lib, 100-day weight. The Ss were maintained within this range throughout the experiment by limited feedings 30 min.

after daily training. During the course of these 10 days, Ss were habituated to the experimental situation by systematic handling, exposure to wet mash pellets scattered on a raised platform, exposure to the runway in groups of three, and finally, two direct placements in the goal box of the alley with a wet mash pellet (.5 gm dry weight) in the food cup.

Experimental training for all Ss consisted of one alley trial per day for 88 days. A trial was begun by placing S in the start box. When S oriented to the start door, the door was raised, allowing S to traverse the alley. On each trial a wet mash pellet, pressed from a mixture of 2 parts (by weight) Purina Lab Chow powder and 1 part water, was available in the food cup. The S was removed from the goal box immediately after eating and returned to its home cage.

The treatment of the two groups differed only in the amount of reward received during the first of three stages of training. Group LSL received 1.0 gm (dry weight) rewards during the first 43 trials (Stage I). Group SSL received .1 gm rewards during the same period. During the next 16 days (Stage II) both groups received .1 gm reward, and during the last 29 days (Stage III), 1.0 gm reward. Thus in the three stages, Group LSL received first large, then small, and finally returned to large reward, whereas Group SSL received small reward during the first two stages and first encountered the larger reward during Stage III.

Results and Discussion

Figure 1 presents mean Total-Response speeds for the two groups during the three stages of training. During Stage I, when Group LSL received 1.0 gm reward as

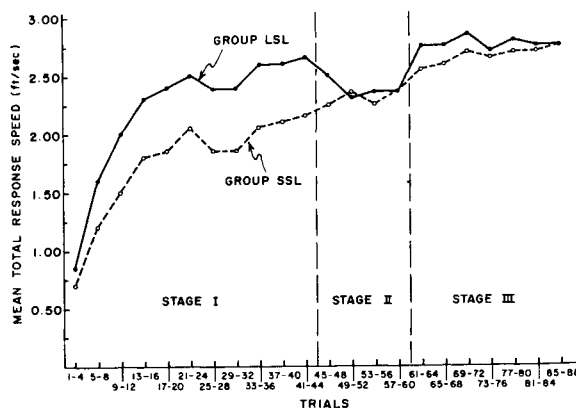


Fig. 1. Mean Total-Response speeds for groups of 20-Ss receiving a sequence of either large, small, large (LSL) or small, small, large (SSL) in three successive stages of runway training. Changed reward size when appropriate according to these sequences was introduced at the termination of trials 44 and 60, after the completion of the response measure reported.

compared to .1 gm for Group SSL, the former group consistently ran faster than the latter. The difference in mean Total-Response speeds on the last four trials in Stage I was highly reliable, $t(38) = 3.40$, $p < .001$.

The response speed of the two groups equalized after 4 trials of Stage II, when both received the same small reward. The speeds remained similar until after the beginning of Stage III, when 1.0 gm rewards were resumed for Group LSL and introduced for the first time for Group SSL. Although both groups received the same magnitude of reward during this stage, the initiation of Stage III was quickly followed by a re-separation of the performance of two groups.

On the first block of 4 Stage III trials following the reintroduction of large reward (i.e. trials 61-64) Group LSL attained a mean speed of running which was maintained with but small fluctuations over the remainder of Stage III. In comparison, Group SSL showed a more gradual increase in running speed, before reaching the same level as Group LSL. A comparison of the increase in running speed on trials 61-64 over the preceding block of 4 trials prior to the receipt of the larger reward yielded a significantly greater change for Group LSL than for Group SSL, $t(38) = 2.43$, $p < .02$.

While the Total-Response measure presents a clear picture of *savings*, Group LSL returning faster to a level of performance appropriate to the larger rewards than Group SSL, this effect was not universally observed in the several component measures. Starting speeds revealed no differences between the groups in Stage III, and goal box speeds showed the largest separation of the two groups. That is, the effect became more pronounced as Ss behavior was measured nearer to the goal area.

According to Hull-Spence theory (e.g., Spence, 1956) different magnitudes of reward may be presumed to lead to the conditioning of anticipatory reward responses ($r_g - s_g$) of different vigors. When Ss are shifted from large to small rewards it may be assumed that a new and less vigorous r_g becomes prepotent, similar to that which exists for consistently small rewarded Ss. The savings effect presently observed suggests that this change is not accompanied by an "unlearning" of the more vigorous r_g . The more vigorous r_g apparently remains available to be more quickly reinstated as prepotent when reward magnitude is again increased.

Another interpretation is possible within the general notion that relearning is faster than original learning. According to Logan's micromolar extension (e.g., 1960) of Hull-Spence theory, it might be assumed that Group

LSL had learned during Stage I to run at a fast speed, and that the savings observed in this group in Stage III was a result of fast responses being more available to again become prepotent.

It could also have been possible that different cues were present on large reward than on small reward trials, so that the superior performance of Group LSL in Stage III could have reflected prior acquisition training to a set of cues to which Group SSL was naive. Common precautions were taken to avoid this possibility, e.g., Ss could not see the reward until after the response measures were taken, and a 24-hour intertrial-interval was employed to minimize aftereffect cues from the preceding reward. One empirical check on the absence of differential cues other than reward aftereffects, is the performance of the two groups on trial 60, on which reward magnitude was shifted from small to large for all Ss, but on which the change should not have been appreciated by S until after the response measures were completed. On this trial there were no differences between the two groups on any of the response measures and no significant changes from trial 59 in either group.

As pointed out, the present data may be consistently integrated within Hull-Spence theory. This theory assumes that reward magnitude does not influence the habit strength (H) of the running response, but rather effects the performance of the running response through the incentive motivational properties of $r_g - s_g$. The data do not, however, agree with the casual generalization that reward magnitude influences "performance" but not "learning." Indeed the present results suggest that the history of reward magnitude may have relatively persistent effects, indicative of "learning" although it may not influence "performance" in all situations.

References

- Crespi, L. P. Amount of reinforcement and level of performance. *Psychol. Rev.*, 1944, 51, 341-357.
- Hulse, S. H. Amount and percentage of reinforcement and duration of goal confinement in conditioning and extinction. *J. exp. Psychol.*, 1958, 56, 48-57.
- Logan, F. A. *Incentive*. New Haven: Yale University Press, 1960.
- Spence, K. W. *Behavior theory and conditioning*. New Haven: Yale University Press, 1956.
- Wagner, A. R. Effects of amount and percentage of reinforcement and number of acquisition trials on conditioning and extinction. *J. exp. Psychol.*, 1961, 62, 234-242.
- Zeaman, D. Response latency as a function of amount of reinforcement. *J. exp. Psychol.*, 1949, 39, 466-483.

Note

1. This research was supported in part by National Science Foundation Grant No. G-24015. The authors wish to thank Mr. Peter Frey for his assistance.