Reinforcement magnitude effects with overtraining

LOIS REEL HAMMER

The George Washington University, Washington, D.C. 20006

A standard runway experiment on shifts in reinforcement magnitude was repeated, but with a much more extended acquisition phase. After 50 trials, the group receiving large and small reward did not differ. Subsequent increase or decrease in reward did not change running times appreciably. Because much competing behavior was observed in this situation, the appropriateness of runway measures is questioned.

Despite a voluminous literature on the effects of reinforcement magnitude, one finds that results are seldom replicated exactly. Most experiments have been carried out in a straight runway using a discrete-trials procedure. In this situation, the main effects are fairly consistent. More disagreement seems to arise over other effects, e.g., whether so-called contrast effects that sometimes occur after shifts in magnitude ("elation" and "depression") are real or artifactual.

It is notable that operant techniques using free response rate as the dependent variable are seldom as successful in showing variations in rate as a function of magnitude (cf. Neuringer's 1967 discussion). Error scores for discrimination learning (Lawson, 1957) or maze learning (Maher & Wickens, 1954) also may not be affected by size of reward.

Rather than conclude that magnitude of reward has so little generality, a comparison of experimental techniques may suggest other reasons why the effects of this variable are hard to demonstrate consistently. It may be that the various dimensions or kinds of behavior that are measured are not influenced equally by reinforcement parameters. Another possibility is that the effect of reward varied depending on the strength or extent of conditioning. The traditional runway experiment is limited to the stage of acquisition in which behavior changes most rapidly. A barpressing response, on the other hand, is so quickly acquired that interest is usually centered on maintenance or steady-state behavior.

It has been suggested (Keesey & Kling, 1961) that, after considerable overtraining, shifts in amount of reward have little effect. Surprisingly little attention has been given to this parameter. The few runway experiments in which number of preshift trials was varied are inconsistent in regard to rate and degree of change in running times (Ashida & Birch, 1964; Davis & North, 1968; Rosen, 1966; Spence, 1956;

Vogel, Mikulka, & Spear, 1967). Variations in kind and amount of reinforcement and other procedural differences contribute to the confusion typical of the problem area in general. Hence, it seemed important to test the hypothesis that changes in reinforcement magnitude have less effect with overtraining by following closely the procedure used in a standard and often cited experiment.

The Zeaman (1949) experiment, selected as representative of the early work on reinforcement shifts, was replicated as closely as possible except for number of trials. In addition, the present experiment served as a test of the reliability of the reinforcement magnitude effect itself using a particular method. In the course of the study, incidental observations were made that seem especially germane to the methodology question.

METHOD

Except for number of training trials and a few presumably minor variations noted below, the experimental variables (choice of reinforcement, response measures, etc.) reported by Zeaman were reproduced as closely as possible.

The Ss were eight male Sprague-Dawley albino rats, about 150 days old, randomly assigned to two groups to receive either large or small reward during the acquisition phase. (The scores from one S that died midway in the experiment were eliminated.) The Ss were maintained at 80% ad lib weight throughout the experiment (a departure from Zeaman's 23-h deprivation, consistent with present laboratory practices).

The experiment was conducted in a 3-ft runway patterned after the original except that it was enclosed by a wire mesh cover rather than open and elevated. The startbox and running alley were painted light gray, the goalbox black. A 15-W light hung 18 in. above the goalbox, dimly illuminated the room.

Food reward was placed in a shallow trough at the back of the goalbox so that it could not be seen from the runway when the goalbox door was open. At the start of a training trial the start door was closed and the goal door was open. Fifteen

seconds after S was placed in the box facing the door, a trial was begun by E's opening the start door and at the same time starting an electric timer operated by a hand switch. The timer was stopped when S passed a line 1 ft from the start door. The elapsed time, termed "response time," was the only measure used. As soon as S had entered the goalbox, the goal door was closed.

The rewards were pieces of processed cheese (Velveeta brand), 2.5 or 0.1 g. (The small reward is twice that used by Zeaman, as it was not possible to weigh out a smaller piece accurately. The amounts still differed by a factor of 25, a substantial difference.)

Ss were given one trial a day throughout the experiment at approximately the same time each day. They were returned immediately to their home cages and fed their daily ration of chow from 1 to 2 h later. Pretraining consisted of two sessions of free exploration of the runway 5 min each day, then two sessions with placement in the goalbox, where 0.5 g of cheese was provided. On Day 5 the first acquisition trial was run.

Fifty acquisition trials were run before reinforcement magnitude was shifted. On the 51st trial the amounts were reversed for the two groups. The postshift phase was terminated with the 69th trial. As Zeaman considered his Ss to have reached asymptote by the 19th trial, this represented 150% overtraining with the original reinforcement magnitudes, according to the usual definition. The choice of 50 trials was somewhat arbitrary but was made before any of the results were known.

RESULTS AND DISCUSSION

Looking first at the postshift measures (Fig. 1), it would appear that the hypothesis is supported. Except for some irregularity in the large-small group, the response times changed very little from pre- to postshift phases. The two groups were not significantly different on the last six trials, using the Mann-Whitney U test.

The preshift performance was more striking, however, and is a feature that calls for more comment. During this phase, the large-reward group showed a rapid decline in response times. If the experiment had been terminated after 30 trials, say, this could have been taken as final or asymptotic performance. Yet there was a further steady decrease in response times up to the last few trials. Spence (1956) reported similar findings in support of his argument that an apparent positive contrast effect occurred only because preshift training had not been carried out to asymptote. In the present case, however, even the small-reward group had reached

*The experiment was carried out with the aid of a grant from NIH Biomedical Sciences Support funds to The George Washington University.

Fig. 1. Mean running times for groups trained with large and small reward. At point labeled "reward shift" the magnitudes were reversed.

the same level by the 50th trial (Mann-Whitney U was not significant over the last six trials). Again, if only 30 trials had been run, it would have been concluded that reinforcement magnitude determines asymptote, consistent with earlier interpretations (Pobols, 1960). Instead, it appears that rate of learning, not asymptote, differentiated the two groups. A similar diminishing effect of reinforcement magnitude with continuing trials is found in Wagner (1961), but that author did not draw attention to it.

The performance of the small-reward group was very erratic during most of the preshift phase. The irregularity was due to extremely long response times shown by one or another S on one trial or another. Contrary to the best statistical usage, these extreme scores are included in calculation of the mean rather than some other index just to indicate some of the variability that was present in the raw data. The behavior of these Ss was qualitatively different from that of the large-reward group, as well as from their own behavior later in the experiment, and included a great deal of competing behavior, first described in this context by Pereboom & Crawford (1958). That is, the response latencies include time spent sniffing the corners, turning around, grooming, and engaging in other kinds of behavior, in addition to running. It was the author's observation that almost all of the competing behavior occurred in the startbox rather than in the runway. This fact reflects a particular weakness in the use of latency measures for instrumental

behavior of this sort, although running times can be subject to the same bias (Marx & Brownstein, 1963). The problem is not obviated by E's waiting a fixed time before opening the door and attempting in other ways to make the conditions at the start of the trial constant. There was a definite point beyond which the competing behavior began to drop out. Thereafter all the response times steadily improved. The qualitative change in behavior over the course of the experiment is taken to reflect a difference in rate of acquisition of the running response itself as a function of the smaller reward.

It was expected that the large-small shift would introduce variability in the other group in the form of similar competing behavior. On the contrary, only three of the individual postshift scores include the long response times characteristic of the other group early in acquisition (a mild "depression" effect?). Throughout the last 10 or more trials, the behavior of all Ss was regular, and both groups overlapped considerably.

Admittedly, this method of replication is limited by the lack of a control group run under the original conditions, i.e., shift after 19 trials. Nevertheless, it is concluded that under the conditions reported by Zeaman (1949) and followed scrupulously in the present experiment, (1) differences in amount of reinforcement do *not necessarily* produce different asymptotic response levels, and (2) after extended training, sizable reduction or increase in amount of reinforcement does not produce

immediate, large changes in behavior. The extent to which these effects do occur may depend in large part on the kind of behavior that is measured. The use of a small number of Ss does not invalidate these conclusions but rather highlights the problem. Qualitative differences in behavior cannot be eliminated by averaging over a large N. Such differences may well contribute to the varying results that are reported, yet the problem is rarely mentioned in this context.

Schrier (1967) argues that positive contrast effects cannot be demonstrated in the runway situation. I concur and question further whether this situation is appropriate for studying *any* effects of reinforcement magnitude.

REFERENCES

- ASHIDA, S., & BIRCH, D. The effects of incentive shift as a function of training. *Psychonomic Science*, 1964, 1, 201-202.
- CAMPBELL, B. A., & KRAELING, D. Response strength as a function of drive level during training and extinction. *Journal of Comparative & Physiological Psychology*, 1954, 47, 101-103.
- DAVIS, S. F., & NORTH, A. J. The effect of number of large reward training trials on behavior following incentive reduction. *Psychonomic Science*, 1968, 11, 311-312.
- KEESEY, R. E., & KLING, J. W. Amount of reinforcement and free-operant responding. *Journal of the Experimental Analysis of Behavior*, 1961, 4, 125-132.
- LAWSON, R. Brightness discrimination performance and secondary reward strength as a function of primary reward amount. *Journal of Comparative & Physiological Psychology*, 1957, 50, 35-39.
- MAHER, W. B., & WICKENS, D. D. Effect of

- differential quantity of reward on acquisition and performance of maze habit. *Journal of Comparative & Physiological Psychology*, 1954, 47, 44-46.
- MARX, M. H., & BROWNSTEIN, A. J. Effects of incentive magnitude on running speeds without competing responses in acquisition and extinction. *Journal of Experimental Psychology*, 1963, 65, 182-189.
- NEURINGER, A. J. Effects of reinforcement magnitude on choice and rate of responding. *Journal of the Experimental Analysis of Behavior*, 1967, 10, 417-424.
- PEREBOOM, A. C., & CRAWFORD, B. M. Instrumental and competing behavior as a function of trials and reward magnitude. *Journal of Experimental Psychology*, 1958, 56, 82-85.
- PUBOLS, B. H., JR. Incentive magnitude, learning, and performance in animals. *Psychological Bulletin*, 1960, 57, 89-115.
- ROSEN, A. J. Incentive-shift performance as a function of magnitude and number of rewards. *Journal of Comparative & Physiological Psychology*, 1966, 62, 487-490.
- SCHRIFER, A. M. Effects of an upward shift in amount of reinforcer on runway performance of rats. *Journal of Comparative & Physiological Psychology*, 1967, 64, 490-492.
- SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale University press, 1956.
- VOGEL, J., MIKULKA, P., & SPEAR, N. E. Effect of interpolated extinction and level of training on the "depression effect." *Journal of Experimental Psychology*, 1967, 72, 51-60.
- WAGNER, A. R. Effects of amount and percentage of reinforcement and number of acquisition trials on conditioning and extinction. *Journal of Experimental Psychology*, 1961, 62, 234-242.
- ZEAMAN, D. Response latency as a function of the amount of reinforcement. *Journal of Experimental Psychology*, 1949, 39, 466-483.

Attenuation of stimulus sensitivity by scopolamine

KENNETH BROWN and DAVID M. WARBURTON*
Queen's University, Belfast, Northern Ireland

Five male albino rats were trained on a DRL 15-sec schedule of reinforcement and injected with six doses of scopolamine hydrobromide. Analysis of the results by means of signal-detection theory indicated that the effects were due to modifications of stimulus sensitivity rather than changes in response bias. This was interpreted as evidence against a response-inhibition hypothesis and suggests that there were changes in the stimulus input.

A number of studies have shown that scopolamine hydrobromide, a peripheral cholinergic blocking agent (Goodman & Gilman, 1965) increased responding in situations requiring response suppression (Heise, Laughlin, & Keller, 1970; Warburton & Groves, 1969). Other studies (Warburton, 1969) have demonstrated similar behavioral effects when cholinergic function was modified by anticholinesterases. This disruption has been attributed to either impairment of stimulus control (e.g., Heise et al, 1970) or to response disinhibition (Bignami, 1967). These hypotheses were tested by using a form of analysis based on the theory of signal detection (TSD). This theory originated in work on electronic communication systems and has been extended to sensory psychology (Green & Swets, 1966). It has been particularly useful in psychophysics because it yields independent estimates of a S's sensitivity to stimuli, d' , and his response bias, β .

The present experiment applied the TSD analysis to performance in a situation involving response suppression, the differential

reinforcement of low rate (DRL). In the DRL schedule a hungry animal is reinforced with food only if his responses are spaced at greater than a specific interval, while responses within this time limit reset the interval (Ferster & Skinner, 1957). The nature of the discriminative stimulus controlling the spaced responding is unknown, but the TSD analysis still enabled the relative changes in sensitivity and bias to be examined.

METHOD

The Ss were five adult male albino rats (Wistar strain), housed individually and maintained at 80% of their "free-feeding" weight.

They were trained to leverpress in a Grason-Stadler (Type E 3125 B) operant box using 45-mg Noyes pellets of food. Thereafter, they were reinforced for spacing their responses at intervals increasing progressively from 5 to 15 sec over a period of 6 days. Supplements of 15 g of dry laboratory diet were given after each daily session. They were maintained on the DRL 15 sec for 50 1-h sessions until the behavior was stable. Prolonged training was needed to satisfy the TSD assumption of well-practiced Ss, i.e., Ss that have an ordinal scale of likelihood ratio that a particular event was caused by a signal

rather than noise (Green & Swets, 1966).

After stable performance was established, doses of 0, 0.0625, 0.125, 0.25, 0.5, and 0.75 mg/kg of scopolamine hydrobromide were injected intraperitoneally on alternate days, according to a Latin square design, providing that performance returned to the baseline level on the intervening day. Previous studies had shown that this interval between injections was sufficient (Warburton & Groves, 1969), and our data confirmed this.

RESULTS

The experimental results were recorded in terms of the number of responses, the number of reinforcements, and the interresponse times (IRTs), using 10 3-sec intervals. From the IRT data, estimates were made of the probability of responding in a particular time interval, on the condition that the animal reached the initial boundary of that interval and so had an opportunity for responding during the interval, the IRTs/OPs analysis used by Anger (1963). From these values, the mean probability of a response was calculated for each animal over four intervals of 3 sec, prior to the availability of reinforcement, giving the "false-alarm" rate, $p(S/n)$. The results for the first 3-sec interval were not used in this calculation because they represented deprivation conditions, apparatus, etc. (Harzem, 1969). The mean probability of a response for the five intervals between 16 and 30 sec was computed to give the "hit" rate, $p(S/s)$. It was possible to average these values because of the relative stability of the IRTs/OPs measure within both the pre- and postreinforcement intervals. The false-alarm rates and hit rates of each animal for each dose level were plotted on double probability graph paper, and an index of sensitivity, d'_e , was estimated for each one by the procedure suggested by J. P. Egan and described in Green & Swets (1966). The individual changes in sensitivity for each dose are plotted in Fig. 1.

These dose-response functions show that the sensitivity to the signal decreased inversely with dose in all animals. It should be emphasized that the computation assumes that the underlying noise and signal + noise distributions are both Gaussian and equal, but, even if they were not, the relationships between d'_e and dose level would be unchanged.

A discriminative index of the hit rate over the false-alarm rate was calculated for each dose level for each S. There was no overlap between the indices for the 0- and 0.0625-mg/kg dose levels, and no overlap between these two and the other four dose

*Present address: Department of Psychology, Reading University, Reading, England.