

Conditioned reinforcement based on reinforcing electrical stimulation of the brain: Chain schedules

RICHARD J. BENINGER and PETER M. MILNER
McGill University, Montreal, Quebec, Canada H3A 1B1

A brief critical review of experiments on conditioned reinforcement based on electrical stimulation of the brain (ESB) revealed that the reported failures to show the effect were a result of methodological aspects of the research; conditioned reinforcement based on ESB has been demonstrated by several experimenters and is, therefore, a reliable phenomenon. Consistent with this was the observation in five rats of positively accelerated responding through the fixed interval (FI) of a chain FI 1-min random-interval (RI) 20-sec schedule where the RI component was in effect for 1 min. Reinforcement consisted of signaled trains of ESB. In a control procedure, three rats, trained on a tandem FI 1-min RI 20-sec schedule, showed much higher FI response rates and less positively accelerated responding. It was concluded that conditioned reinforcement was observed in the chain and that traditional theories of the behavioral effects of ESB could not account for these and earlier data.

Brief trains of electrical stimulation of the brain (ESB) can be used to establish and maintain operant responding (Olds & Milner, 1954), thereby making ESB a positive reinforcing stimulus (Skinner, 1938). Typically, neutral stimuli that are associated with a reinforcer can themselves acquire the properties of a reinforcer; for example, Skinner (1938, pp. 82-83) showed that a sound previously associated with the presentation of food to a food-deprived rat could be used to condition a new response. The sound, therefore, was defined as a conditioned reinforcer.

There is some question about the conditions necessary for a neutral stimulus associated with ESB to acquire reinforcing properties. As a result of this uncertainty, it has been suggested that the behavioral effects of ESB and conventional (e.g., food) reinforcing stimuli differ (Gallistel, 1973). It is our contention that conditioned reinforcement based on ESB primary reinforcement is a reliable phenomenon. This paper will review previous experiments showing that failures to observe conditioned reinforcement based on ESB can be attributed to methodological aspects of the research rather than to ESB per se. In addition, data are presented from rats trained to respond for ESB according to chain schedules; these data provide additional evidence that stimuli

associated with ESB do become effective conditioned reinforcers.

Studies of conditioned reinforcement can be classified, on the basis of their methodology, as using one of the following: (1) established response procedure, (2) new response procedure, or (3) chain response procedure. To our knowledge, all of the experiments in the literature which deal with conditioned reinforcement based on ESB fit into the first two classifications.

The prototype for studies using the established response procedure was provided by Bugelski (1938). He trained rats to press a lever for food with each presentation of food preceded by a signal. All rats were then placed on extinction, half continuing to receive the signal for pressing the lever while, for the other half, lever responses had no programmed consequences. The rats receiving the signal made more extinction responses than the no-signal rats, and Bugelski concluded that the signal had become a conditioned reinforcer. A similar procedure was employed by Mogenson (1965) and Seward, Uyeda, and Olds (1959) using ESB in the training phase. In neither of these studies, however, did the extinction responding of the signal and no-signal groups differ, and it was concluded that conditioned reinforcement based on ESB was not attained.

A closer look at the three experiments reveals some differences, however. Bugelski (1938) used two groups of 32 rats *each* and reported a difference in extinction responding with only 95% confidence. This suggests that the interrater variability was high in his experiment and that this variability would probably mask a small effect if the groups were smaller.

The authors wish to thank the following for assistance: S. Amir with surgery and J. Bayreuther in running the animals. This research was supported by National Research Council of Canada Grant AP66 to P.M. Milner. Reprints can be obtained from the authors, Department of Psychology, McGill University, 1205 McGregor, Montreal, Quebec, Canada H3A 1B1.

Since Seward et al. (1959) used a total of 14 rats and Mogenson (1965) used 11, even if there was a conditioned reinforcement effect, their data were not likely to show it. These differences in the number of rats may account for the failures to observe conditioned reinforcement based on ESB using the established response procedure. Quite apart from these considerations, there is considerable doubt about the efficacy of the established response procedure for studying conditioned reinforcement (Mackintosh, 1974, pp. 234-235) and it is no longer used for this purpose.

The new response procedure was first used by Grindley (1929) and subsequently by Skinner (1938). Skinner presented rats with food pellets on one day and then presented just the operation of the pellet dispenser but no pellets for leverpresses during a subsequent session. All rats learned to press the lever, and Skinner concluded that the click of the dispenser was a conditioned reinforcer. A similar method has been used by Knott and Clayton (1966), Mogenson (1965), Stein (1958), and Trowill and Hynek (1970) to study conditioned reinforcement based on ESB. Three of these studies reported that stimuli previously associated with ESB could be used to train a new response; they also showed that the new response could be made discriminantly, eliminating the possibility of attributing the effect to, for example, changes in general activity (cf. Mackintosh, 1974, pp. 235-237).

The only study which failed to show conditioned reinforcement based on ESB using the new response procedure was a replication of Stein (1958) by Mogenson (1965). It appears that, in this study, the onset of the light or buzzer (stimuli to be associated with ESB) was contiguous with the onset of ESB during the conditioning phase. Since stimuli do not become conditioned reinforcers unless they briefly *precede* the primary reinforcer (e.g., by 0.5 sec; Kelleher, 1966, p. 181), Mogenson's experiment would not be expected to show conditioned reinforcement.

From this brief review, it is concluded that ESB can be reliably used to establish conditioned reinforcement. Previous failures to observe the effects are not due to the use of ESB per se. Since it has been shown that animals can be trained to respond on schedules of intermittent reinforcement for ESB (Beninger, Bellisle, & Milner, in press; Cantor, 1971), it should be possible to also use scheduling techniques to demonstrate conditioned reinforcement. In the following experiment, rats were trained to respond on a chain fixed-interval (FI) random-interval (RI) schedule for ESB; i.e., completion of the FI response contingency produced a stimulus associated with the RI schedule and the opportunity to respond for ESB according to the RI schedule.

If the stimulus associated with the RI schedule has acquired reinforcing properties, responding in the FI component should be appropriate to the FI schedule, i.e., positively accelerated through the interval (Ferster & Skinner, 1957).

METHOD

Subjects

Five male Long-Evans hooded rats were individually housed and maintained on free food and water throughout the experiment. Rats 252, 260, 261, and 265 had had 20-22 sessions of training on RI schedules previous to the initiation of this experiment; rat 294 was experimentally naive.

Surgery and Histology

After anesthetization with IP injection of 50 mg/kg Nembutal, each rat was stereotaxically implanted with one Plastic Products bipolar electrode (No. MS 303-0.010) made of intertwined stainless steel wires 0.25 mm in diameter. With the incisor bar set at 3.2 mm above the interaural line, coordinates of 4.0 mm posterior to bregma, 1.5 mm lateral to the midline, and 8.0-8.5 mm below the surface of the skull were used for rats 252 and 265, with the substantia nigra as the target; for the remaining rats, respective coordinates were 0.8, 1.5, and 8.5-8.6 mm, with the lateral hypothalamus as the target. At least 6 days of recovery preceded initiation of the experiment.

At the completion of the experiment, the rats were killed and perfused through the heart with 10% Formalin. The brains were extracted, fixed in 10% Formalin, sliced at 40 μ m and stained with thionin to verify electrode loci.

Apparatus

The experimental environment consisted of a Plexiglas cubicle (approximately 22 cm³) with a grid floor and one aluminum plate wall on which was located a lever (1.8 cm wide; Lehigh Valley Model 121-03) with its surface at a height of 2.0 cm. The force requirement for the lever was 10.2 g. The cubicle was located in a ventilated sound-attenuating box, which was equipped with a speaker and a relay which could be used as noise and click sources, respectively.

ESB was provided through leads attached to a mercury commutator, thereby allowing the rats freedom of movement. Electrical stimulation consisted of 0.5-sec trains of rectangular biphasic pulses, 0.1 msec in duration and presented at 100 Hz. The current varied from rat to rat, ranged from 200 to 400 μ A, and was kept constant by placing a large resistance (10⁵ ohms) in series with the rat.

BRS/LVE solid-state switching and timing devices were used to arrange the contingencies. RI schedules were generated by a probability gate (BRS Electronics Model PP-1) pulsed by a variable time source. Data were recorded on electromechanical counters and a timer.

Procedure

For rat 294, the leverpress response was shaped with reinforcement consisting of a train of ESB, the onset of which was contiguous with the response. Shaping was followed by two 30-min daily sessions of training on a RI 30-sec schedule; i.e., a leverpress response was reinforced after an average interval of 30 sec. For all subsequent sessions, reinforcement consisted of a train of ESB which was signaled by a click 0.5 sec before its onset. Five sessions of RI 30-sec, during which white noise was continuously present, were followed by three sessions during which a chain FI 30-sec RI 20-sec schedule was in effect. During the FI schedule, the white noise was off; completion of the FI contingency (i.e., the first response after 30 sec) turned the white noise on and switched the RI schedule into effect for 30 sec. After 30 sec, the

noise went off automatically and the FI was again in effect. Following these training sessions, the schedule was changed to chain FI 1-min RI 20-sec. The RI schedule was associated with white noise and remained in effect for 1 min, after which the RI was automatically switched to the FI schedule and the noise turned off. A session begun with the RI component and was usually terminated after 20.5 cycles of the chain, a minimum of 41 min.

Rats 261 and 265 had experienced leverpressing for ESB in the presence of white noise and were placed directly onto the chain FI 1-min RI 20-sec schedule as described above. Training on this schedule lasted 20-30 sessions, as indicated for each rat on the x-axis of Figure 1.

As a control procedure, experienced rats 252 and 260 were placed directly on a tandem (tand) FI 1-min RI 20-sec schedule. This schedule operated in the same manner as the chain (cf. Ferster & Skinner, 1957), except that the noise was always on. Training on the tand schedule continued for 35 sessions, usually with 20.5 cycles per session.

For rat 294, the chain schedule training was followed by 20 sessions of the tand schedule; for rats 252 and 260, the tand schedule training preceded 23 sessions on the chain schedule. Rats 261 and 265 died after 24 and 30 sessions, respectively, on the chain.

Dependent variables were latency to the first response in each session and response rate (responses per minute). FI response rates were accumulated on three counters representing three 20-sec segments of the 1-min interval.

RESULTS

Histology

The brains of three rats were lost, two because the rats died overnight and one because the brain was badly damaged during sectioning. For rat 252, the electrode tip was in the dopaminergic A9 cell group (Dahlstrom & Fuxe, 1964) of the substantia nigra zona compacta; for rat 294, the electrode tip was on the ventral border of the zona incerta dorsal to the fornix and probably stimulated catecholamine fibers which pass just ventral to it (Jacobowitz & Palkovits, 1974).

Behavior

Response rates in responses per minute for the rats which had the chain condition first are shown in Figure 1. The histograms consist of sets of three bars per session, showing the response rates for the first, middle, and last segments of the FI. The broken portion of the third bar of each set indicates the contribution that changeover responses make to the rate for that segment; changeover responses are the responses that switch the RI schedule into effect. Strictly speaking, these responses occur outside the third segment of the FI, since the FI must elapse before a changeover can occur. They are, however, traditionally included in the terminal segment of the FI. The graphs in Figure 1 show RI response rates for each session. Only the last 15 chain or tand sessions are shown for each rat.

During the chain FI 1-min RI 20-sec schedule, rats 261, 265, and 294 clearly discriminated the two

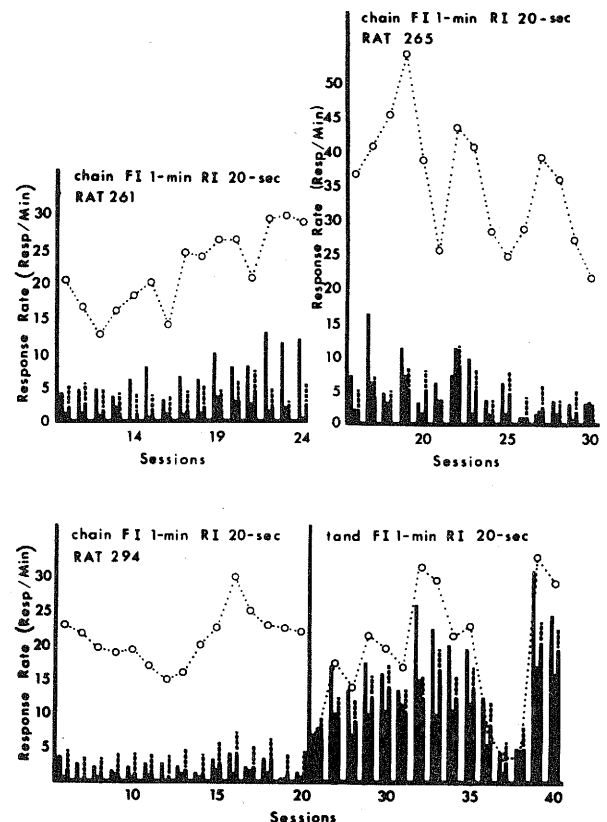


Figure 1. Response rates (responses/minute) for the rats which had the chain condition first. Histograms consist of sets of three bars per session showing the rates for the first, middle, and last segments of the FI. The broken portion of the third bar of each set indicates the contribution that changeover responses make to the rate for that segment (see text). The graphs show RI rates.

components, response rates during RI being consistently much higher than during FI. Within sessions, the pattern of FI rates was almost always U-shaped; i.e., responding was higher at the beginning of the interval, probably because of runover from the much higher RI rates, low in the middle of the interval, and high again at the end.

Only rat 294 received the tand condition after the chain. RI rates for rat 294 were often still slightly above FI rates but were not nearly as different from FI rates as they were during the chain condition. FI rates were much higher during the tand than during the chain, but still often showed the U-shaped pattern seen in the FI of the chain. However, the rate of FI responding during the last, as compared to the middle, segment was usually much higher in the chain than in the tand condition, and if the rate in the last segment of the FI is expressed as a ratio of the rate in the middle segment the means of these ratios differ significantly according to a *t* test, $t(14) = 4.72, p < .001$.

Figure 2 shows the response rates for rats 252 and

260, the rats which received the tand condition first. During the tand condition, RI rates for both rats were slightly higher than rates in all segments of the FI for about half of the sessions, while in the other half, FI and RI values were similar. FI response patterns were often U-shaped but, as noted above, the rates of responding in the middle and last segments were more similar in the tand than in the chain. This can be shown to be statistically true by doing a *t* test for each rat, comparing the means of the ratios of the response rates in the third to second segments of the FI: $t(14) = 6.05$, $p < .001$, for rat 260 and $t(14) = 2.56$, $p < .025$, for rat 252. RI rates in the chain were consistently much higher than FI rates, while FI patterns for rat 252 were U-shaped. Rat 260 seldom responded during the FI at all, but continued to turn on the white noise, usually after longer than a 1-min interval, and responded during the RI throughout the 23 chain sessions.

DISCUSSION

Was the white noise a conditioned reinforcer for responding during the FI component of the chain schedule? The necessary conditions for answering this question affirmatively would be the observation of responding appropriate to the FI in that component of the chain; in the FI component of the tand schedule, responding should be similar to responding in the RI component, since no discriminative stimuli are presented and reinforcement becomes available at random intervals.

In dealing first with the latter condition, response rates in the FI and RI components of the tand (observed in three rats) were quite similar. In the 20-sec segments of the FI component of the tand, however,

the response pattern was observed to be U-shaped. This observation suggests that the rats were making a temporal discrimination. Since both the RI and FI schedules remained in effect for 1 min, rats responding on the tand may have learned to discriminate periods of high and low reinforcement density. This position is supported by the observation that ESB can act as a discriminative stimulus (Butcher & Stutz, 1969; Stutz & Asdourian, 1965; Stutz, Butcher, & Rossi, 1969). The discrimination would result in a cycle of responding, with peak responding during the RI and lowest responding during the middle 20-sec segment of the FI, exactly the pattern that was observed in the tand for all rats.

Under the chain condition, all rats showed a clear discrimination of components, responding at much higher rates during RI than during FI. Within the FI of the chain, the rats showed a U-shaped pattern of responding; if the high rate in the first 20-sec segment of the FI was caused by runover from the RI, a phenomenon often observed in animals responding on FI schedules for conventional reinforcers, then this pattern of responding is the typical FI scallop (i.e., positively accelerated responding through the interval) reported by Ferster and Skinner (1957). Since the amount of scalloping was seen to be significantly higher in the FI of the chain than in the FI of the tand and since this pattern of responding was observed in a FI where only the onset of a stimulus (noise) associated with primary reinforcement was the outcome, that stimulus was a conditioned reinforcer (Skinner, 1938).

The data presented here show that conditioned reinforcement based on ESB can be observed using chain schedules and are consistent with the data of Knott and Clayton (1966), Stein (1958), and Trowill and Hynek (1970). In addition, these data supplement and extend those of Beninger et al. (in press) and Cantor (1971) which showed that signaled ESB was capable of maintaining responding on schedules of intermittent reinforcement.

The observation both of consistent responding for intermittent ESB and of reliable conditioned reinforcement based on ESB poses considerable problems for the theory of ESB effects put forward by Deutsch and Howarth (1963) and by Gallistel (1973). This theory states that the presentation of an ESB activates reinforcement and "drive" systems in the brain and that the strength of this activation decreases to a low level in the minute or two following ESB. As the ESB-produced "drive" decays, so does the probability of further responding.

According to this theory, responding for intermittent ESB is not attainable because the necessary "drive" for the responding has time to decay. In the present experiment, as the 1-min interval of the FI passed, the "drive" for further responding should have become weaker and weaker, especially because

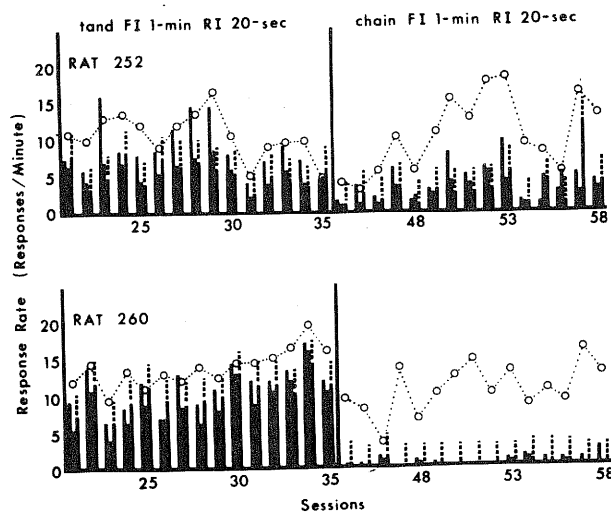


Figure 2. Response rate (responses/minute) for the rats which had the tand condition first. See the caption of Figure 1 for description.

our rats had no apparent biological need (i.e., they had ad-lib food and water), but instead responding was observed to *increase* with time since the last ESB, showing the traditional FI scallop. Moreover, the latency to initiate responding at the beginning of a session, for which the "drive," according to the Deutsch theory, should be at a low level because at least 23 h has elapsed since the last ESB, was observed to be less than 15 sec in 74% of the sessions shown in Figures 1 and 2 and in only 7% of the sessions was the latency greater than 60 sec. These data suggest that ESB-produced "drive" is not necessary for responding for ESB to occur.

It seems that traditional theories of ESB, and more generally of reinforcement, were based partly on the belief that conditioned reinforcement based on ESB was not reliably attainable and that intermittent ESB was ineffective in maintaining behavior. The data presented here and in other papers (Beninger et al., in press; Beninger & Milner, in press) suggest that a restructuring of these ideas for explaining behavior is required.

REFERENCES

- BENINGER, R. J., BELLISLE, F., & MILNER, P. M. Schedule control of behavior reinforced by electrical stimulation of the brain. *Science*, in press.
- BENINGER, R. J., & MILNER, P. M. The effects of signalled and unsignalled brain stimulation, water and sucrose reinforcement on running behavior in rats. *Journal of Comparative and Physiological Psychology*, in press.
- BUGELSKI, R. Extinction with and without sub-goal reinforcement. *Journal of Comparative Psychology*, 1938, **26**, 121-134.
- BUTCHER, R. E., & STUTZ, R. M. Discriminability of rewarding subcortical brain shock. *Physiology and Behavior*, 1969, **4**, 885-887.
- CANTOR, M. B. Signalled reinforcing brain stimulation facilitates operant behavior under schedules of intermittent reinforcement. *Science*, 1971, **174**, 610-613.
- DAHLSTROM, A., & FUXE, K. Evidence for the existence of monoamine-containing neurons in the central nervous system. I. Demonstration of monoamines in the cell bodies of brain stem neurons. *Acta Physiologica Scandinavica*, 1964, **62**, Suppl. 232, 1-55.
- DEUTSCH, J. A., & HOWARTH, C. I. Some tests of a theory of intracranial self-stimulation. *Psychological Review*, 1963, **70**, 444-460.
- FERSTER, C. B., & SKINNER, B. F. *Schedules of reinforcement*. New York: Appleton-Century-Crofts, 1957.
- GALLISTEL, C. R. Self-stimulation: The neurophysiology of reward and motivation. In J. A. Deutsch (Ed.), *The physiological basis of memory*. New York: Academic Press, 1973.
- GRINDLEY, G. C. Experiments on the influence of amount of reward on learning in young chickens. *British Journal of Psychology*, 1929, **20**, 173-180.
- JACOBOWITZ, D. M., & PALKOVITS, M. Topographic atlas of catecholamine and acetylcholinesterase-containing neurons in the rat brain. I. Forebrain (telencephalon, diencephalon). *Journal of Comparative Neurology*, 1974, **157**, 13-28.
- KELLEHER, R. T. Chaining and conditioned reinforcement. In W. K. Honig (Ed.), *Operant conditioning: Areas of research and application*. New York: Appleton-Century-Crofts, 1966. Pp. 160-212.
- KNOTT, P. Q., & CLAYTON, K. N. Durable secondary reinforcement using brain stimulation as the primary reinforcement. *Journal of Comparative and Physiological Psychology*, 1966, **61**, 151-153.
- MACKINTOSH, N. J. *The psychology of animal learning*. London: Academic Press, 1974.
- MOGENSEN, G. J. An attempt to establish secondary reinforcement with rewarding brain stimulation. *Psychological Reports*, 1965, **16**, 163-167.
- OLDS, J., & MILNER, P. M. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, 1954, **57**, 419-427.
- SEWARD, J. P., UYEDA, A., & OLDS, J. Resistance to extinction following cranial self-stimulation. *Journal of Comparative and Physiological Psychology*, 1959, **52**, 294-299.
- SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century-Crofts, 1938.
- STEIN, L. Secondary reinforcement established with subcortical stimulation. *Science*, 1958, **127**, 466-467.
- STUTZ, R. M., & ASDOURIAN, D. Positively reinforcing brain shock as a CS in the acquisition and extinction of a shuttle box avoidance response. *Psychonomic Science*, 1965, **3**, 191-192.
- STUTZ, R. M., BUTCHER, R. E., & ROSSI, R. Stimulus properties of reinforcing brain shock. *Science*, 1969, **163**, 1081-1082.
- TROWILL, J. A., & HYNEK, K. Secondary reinforcement based on primary brain stimulation reward. *Psychological Reports*, 1970, **27**, 715-718.

(Received for publication February 10, 1977;
accepted March 17, 1977.)