INVESTIGATIVE ACTIVITY DEPENDING ON REINFORCEMENT CONDITIONS AND INFORMATION INTERACTION WITH THE ENVIRONMENT

A. B. Saltykov, A. V. Toloknov, and N. K. Khitrov

UDC 612.821.1/.3.19].08:591.51

KEY WORDS: learning; investigative reactions; reinforcement conditions; rats.

Investigative activity of animals during learning in chance environments is usually described in terms of two most widely accepted theories. According to one of them, the frequency of instrumental responses is proportional to the probability of the unconditioned reinforcement used [1, 11, 14]. For a multicomponent learning regime this dependence has been generalized in the form of Herrnstein's law [6, 9]. However, it is more convenient to interpret some of the experimental data, which do not conform to Herrnstein's law, from the standpoint of the optimization theory, according to which, behavior is aimed exclusively at optimizing the quantity of reinforcement received [12, 13]. The presence of these different approaches can be conjecturally explained by a difference in the experimental models used and by the possible effect of certain factors not taken into account [6, 10]. One such factor may be the probability of random correct performance of the instrumental reaction (CPIR), which determines what part of the investigative reactions in the initial stage of learning will be undertaken in connection with the conditioned stimulus and appropriately reinforced. This, in turn, may significantly affect the process of information interaction between animal and external environment, modulating its investigative activity.

The existence of optimal values of CPIR, least favorable for learning was demonstrated previously [3, 4]. Meanwhile the mechanisms of this phenomenon have not been adequately studied: only the influence of CPIR on the dynamics of investigative activity has been established (and expressed as normalized relative values) under 100% reinforcement conditions [5].

The aim of the present investigation was to study correlation between the values of CPIR and the absolute values of investigative activity under conditions of 25% and 100% reinforcement

EXPERIMENTAL METHOD

Experiments were carried out on 64 noninbred male rats weighing 250-310 g. The animals were divided into eight groups with eight in each group. An instrumental defensive reflex of pressing a pedal under the influence of a conditioned photic stimulus was produced by the method described previously [3]. The unconditioned stimulus consisted of electrodermal stimulation of the limbs through a grid floor. The rats could escape routine electrodermal stimulation by correctly performing the instrumental reaction (pressing a pedal during photic stimulation). Under these circumstances, in groups 1-4, electrodermal stimulation was withheld in 100% of cases, whereas in groups 5-8, 25% reinforcement was used. The duration of conditioned stimulation and of subsequent intertrial intervals was assigned as uniformly distributed random values, taken from a table. In groups 1 and 5 they were equal to one another; in groups 2 and 6 each intertrial interval was 3 times longer than the previous conditioned stimulus, in groups 3 and 7 it was 7 times longer, and in groups 4 and 8 it was 19 times longer. When the distribution of investigative instrumental reactions was uniform in time (until the time when correlation was found between the conditioned stimulus and reinforcement), mathematical expectation of probability of random pressing on the pedal, against the background of lighting of a small lamp for the above-mentioned pairs of groups (Table 1) was

I. M. Sechenov Moscow Medical Academy. (Presented by Academician of the Academy of Medical Sciences of the USSR K. V. Sudakov.) Translated from Byulleten Éksperimental'nol Biologii i Meditsiny, Vol. 112, No. 11, pp. 451-453, November, 1991. Original article submitted February 20, 1991.

0007-4888/91/0011-1529$12.50

TABLE 1. Conditions of Formation of Instrumental Reflex (limits of variations given in parentheses)

Group No.	Mathematical expectation of PRCR	Duration of conditioned stimulus, sec	Duration of intertrial intervals, sec	Frequency of reinforcement of correct responses, %
1	0,5	20 (1—39)	20 (1—39)	100
2	0,25	10 (1—19)	30 (3—57)	100
3	0,125	5 (1—9)	35 (7—63)	100
4	0,05	2 (1—3)	38 (19—57)	100
5	0,5	20 (1—39)	20 (1—39)	25
6	0,25	10 (1—19)	30 (3—57)	25
7	0,125	5 (1—9)	35 (7—63)	25
8	0,05	2 (1—3)	38 (19—57)	25

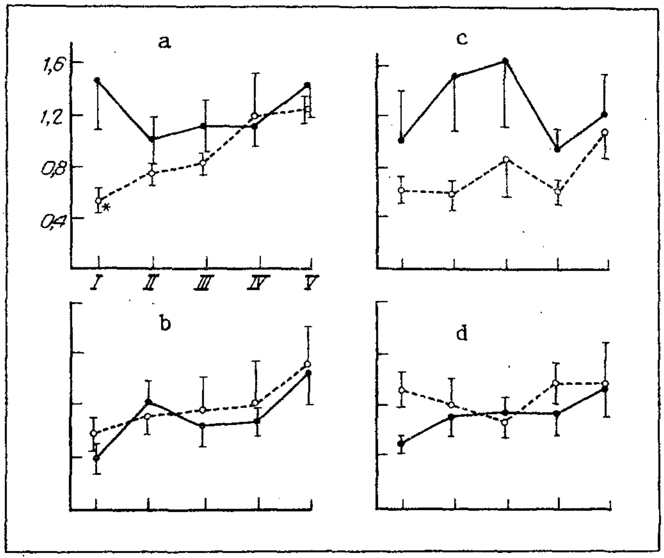

Fig. 1. Dynamics of investigative activity during different learning periods for animals with different values of CPIR and reinforcement conditions. a-d) Groups: a) 1 and 5, b) 2 and 6, c) 4 and 8, d) 3 and 7 (calculated value of CPIR was 0.5, 0.25, 0.05, and 0.125 respectively). Abscissa, learning periods (I-V); ordinate, number of instrumental reactions on average per cycle. Asterisk indicates statistically significant difference between paired groups ($p < 0.05$).

20 sec/(20 sec + 20 sec) = 0.5 sec, 10 sec/(10 sec + 30 sec) = 0.25, etc., respectively.

Investigative activity was assessed by the number of instrumental responses included on average in each cycle (in periods of conditioned and unconditioned stimulation, plus the subsequent intertrial interval). The conditioned reflex was considered to have been formed if the number of correct responses was statistically significantly ($p \leq 0.05$) greater than the a priori level of their random performance [2].

EXPERIMENTAL RESULTS

For graphic representation of the dynamics of investigative activity during learning, five learning periods were conventionally distinguished: the total number of cycles used by each animal to form a conditioned reflex was divided into five equal parts, after which the number of instrumental responses corresponding to each cycle of the corresponding period was counted. A paired comparison was made only of those experimental groups which were trained using equal durations of presentation of the conditioned stimulus and of the intertrial intervals (Fig. 1). All the remaining combinations of possible paired comparisons of the experimental groups (1 – 2, 1 – 3, 1 – 4, 2 – 3, and so on), in our view, were uninformative as regards explaining the effect of the value of CPIR on the dynamics of investigative activity with different frequencies of reinforcement, for under these conditions the action of additional factors had to be taken into account: the absolute duration of presentation of the conditioned stimulus and of the intertrial interval (Table 1). At the same time, we know that a change in each of the above values itself has a marked and often a similar effect on the learning process [6, 7].

As Fig. 1 shows, a change in reinforcement conditions differed in its effect on investigative activity of the animals, depending on the value of CPIR. For instance, if CPIR was 0.5 (groups 1 and 5) a fourfold decrease in the frequency of reinforcement led to a significant reduction of investigative activity throughout the periods of learning (except the 4th), and this was particularly marked in the initial period ($p < 0.05$). A similar result was obtained also when CPIR was 0.05 (groups 4 and 8), which can be taken to be experimental confirmation of the position adopted by several workers on the existence of a direct relationship between the frequency of instrumental responses and the probability of the reinforcement used [8, 9, 11, 14]. This general rule is evidently based on the sharp reduction of information value of the investigative reactions (and, correspondingly, the uselessness of their partial performance) in the case of 25% reinforcement conditions. A number of conclusions drawn previously [4], according to which values of CPIR of 0.05 and 0.5 are least favorable for learning under the conditions of the experimental model used, and they also lead to a decrease in the information value of the investigative reactions, must be particularly emphasized.

If CPIR was 0.25 or 0.125 a qualitatively different rule was observed. Comparison of the 2nd-6th and 3rd-7th paired groups showed that variation of the reinforcement conditions did not lead to consistent and significant changes in investigative activity. During some learning periods a rather higher level of investigative activity was observed under 100% reinforcement conditions, whereas in other cases it was only at 25% (Fig. 1). Under these conditions, the direct relationship between the frequency of the investigative reactions and the probability of unconditioned reinforcement evidently disappears or becomes much weaker. It is a particularly interesting fact that values of CPIR corresponding to this state (0.25 and 0.125) are optimal for learning and weaken the relationship between the rate of reflex formation and the regime of unconditioned reinforcement [3, 4]. With values of CPIR optimal for learning, the information value of each investigative reaction is increased to such an extent that this leads to diminution of the unfavorable effect of the stochastic character of reinforcement on investigative activity.

Thus during learning in random environments investigative activity of animals is determined not only by the frequency of unconditioned reinforcement, but also by the value of CPIR. Which part of the instrumental reactions is performed initially in connection with the conditioned stimulus, and which part will be reinforced appropriately, depend on the values of CPIR. This, in turn, affects the information interaction between animal and environment whatever the conditions of reinforcement (continuous or stochastic). At values of CPIR least favorable for learning reduction of the frequency of reinforcement causes a decrease in investigative activity as a result of sudden worsening of the conditions for receiving information: both factors examined above (CPIR and the stochastic character of reinforcement) lead to a decrease in the informative value of the investigative reactions. Conversely, values of CPIR which are optimal for learning lead to an increase in the informative value of every investigative reaction, and this weakens the relationship between investigative activity and the probability of unconditioned reinforcement.

LITERATURE CITED

1. D. N. Menitskii, Zh. Vyssh. Nerv. Deyat., **33**, No. 5, 803 (1983).
2. A. B. Saltykov, I. V. Smirnov, V. P. Starshov, and M. M. Saltykova, Zh. Vyssh. Nerv Deyat., **36**, No. 5, 987 (1986).
3. A. B. Saltykov, A. V. Toloknov, and N. K. Khitrov, Zh. Vyssh. Nerv. Deyat., **39**, No. 4, 654 (1989).
4. A. B. Saltykov, A. V. Toloknov, and N. K. Khitrov, Zh. Vyssh. Nerv. Deyat., **40**, No. 3, 467 (1990).
5. A. B. Saltykov, A. V. Toloknov, and N. K. Khitrov, Byull. Éksp. Biol. Med., **60**, No. 10, 344 (1990).
6. W. M. Baum, J. Exp. Analys. Behav., **36**, No. 3, 387 (1981).
7. C. A. Bruner, Rev. Mex. Anal. Conducta, **7**, No. 2, 149 (1981).
8. M. Davison and L. Charman, J. Exp. Analys. Behav., **47**, No. 1, 29 (1987).
9. A. Houston, J. Exp. Analys Behav., **45**, No. 1, 15 (1986).
10. H. Imada and Y. Nageishi, Psychol. Bull., **91**, No. 3, 573 (1982).
11. J. Kopp, G. Bourland, R. D. Tarte, and C. R. Vernon, Psychol. Rec., **26**, No. 1, 49 (1983).
12. J. M. Namara and A. I. Houston, Behav. Anal. Lett., **3**, No. 3, 157 (1983).
13. D. A. Stubbs, L. R. Dreyfus, J. G. Fetterman, and L. G. Dorman, J. Exp. Analys. Behav., **45**, No. 3, 333 (1986).
14. W. Vaugham and H. L. Miller, J. Exp. Analys. Behav., **42**, No. 2, 337 (1984).