PARTIAL REINFORCEMENT INCREASES RESISTANCE TO EXTINCTION OF OPERANT TASK IN THE HOUSE MICE *MUS MUSCULUS* TRAINED WITH A CLICKER

Agata Kiełtyka-Kurc, Natalia Solarska, Marcin T. Górecki*

Department of Zoology, Poznań University of Life Sciences, Wojska Polskiego 71C, 60-625 Poznań, Poland
*Corresponding author: Marcin T. Górecki, tel.: +48 61 848 7224,
e-mail address: marcin.gorecki@up.poznan.pl*

Abstract.
Learning skills are crucial in terms of animal chance of survival in nature. However, control experiments in natural conditions are not very effective, then results of training of kept animals and studies on their learning can enlarge our understanding of their skills. The clicker training is a popular form of training of many animal species, consisting of associating a neutral stimulus (a click) with a reward. In our paper we examined if the used reinforcement scheme has an influence on the extinction of learned behaviour. Forty female house mice Mus musculus *were divided into 2 groups of 20 subjects towards which two separate rewarding schemes were used. For completing a learned activity, the mice from the A group received a reward after each click (continuous reinforcement), and the mice from the B group only after the second or the third click (partial reinforcement). Afterwards, at the extinction phase the mice were receiving none reinforcement. Our results showed that the mice which were given only partial reinforcement were more resistant to the extinction of learned behaviour, i. e. performed significantly more attempts to obtain a reward. We believe that our findings could be also considered in natural conditions, where animals could obtain rewards with different frequency in relation e.g. with stable/unstable access to food in a given place.*

Key words: learning; mouse; Mus musculus*; operant conditioning; rodent*

INTRODUCTION

Learning skills are very important in terms of animal chance of survival, e.g. exploitation of novel food, as well as avoiding possible threats such as poisonous and venomous organisms and predators. Results of training of kept animals and studies on their learning can enlarge our understanding of their skills. During the last years using the operant conditioning in animal training became very popular. It is used to train domesticated and wild animals (e.g. Johnson 2003, Dickmann *et al.* 2022), and even snails (Sangha *et al.* 2002) or octopuses (Crancher *et al.* 1972). One kind of operant conditioning training is the clicker training. It consists of associating the click (or other sound signal) with a reward. An animal learns that after the sound signal the reward awaits it - the click becomes a positive reinforcer. A significant advan-tage of this training is the fact that it is not oppres-sive. Another factor convincing trainers to the clicker training is the possibility of immediate reinforcement of trained behaviour. In this case the click is an an-nouncement of a reward. According to Reid (1996), both rewarding and punishing dogs should occur in-stantly after performing the (wanted or unwanted) behaviour. Giving a reward or punishment too late

can delay the already learned behaviour and can cause confusion (Ohnishi *et al.* 2003; 2004). Yama-moto *et al.* (2009) showed that even 0.5 sec of delay in giving commands, rewarding and punishing by the owner decreased the dog's reaction for learned commands. The important element affecting the clicker training is not only the type of the reward and timing but also the frequency of rewarding. There are two basic reinforcement schemes: continuous and partial reinforcement. The continuous reinforcement consists of rewarding (after the click) every time an animal performs a certain activity. The partial reinforcement consists of 'dispersion' of rewarding after the click, so that giving the reward e.g. after the second or third click (Pryor, 2002). Depending on the used reinforcement scheme, the behaviour can be learned in a shorter or longer period of time and be less or more vulnerable to extinction. In this paper we used both reinforcement schemes in order to check if they affect the length of the period of extinction of the learned behaviour in a house mouse. Our study was conducted in captivity, but animals living in natural conditions obtain rewards (reinforcements) with different frequency in relation with e.g. stability of food presence in a given location. Thus our results could

help to understand learning in general, and learning influences all major ecological and evolutionary processes in animals (Dukas 2009).

Materials and Methods

A total number of 40 female albino house mice *Mus musculus* were observed. The experiment was divided into 2 rounds with 20 specimens in each one. All mice originated from a single breeder and were ca 2 months old. The animals were kept in 2 groups of 10 specimens in two glass aquariums with the dimensions of 70cm x 30cm x 45cm. The litter consisted of sawdust and hay. The aquarium was equipped with: clay houses, cardboard rolls (changed every few days), dry branches and little bowls for the water and food. The mice were fed with a complete feed supplemented with fruits, vegetables and dry bread. In order to conduct the experiment individually on every single mouse, two smaller aquariums were used: made of glass (45 x 25 x 25cm) and plastic (45 x 25 x 20cm), equipped similarly to the group aquariums. In each small aquarium one randomly chosen mouse was put and left for 20-24h in order to familiarise with the environment. For the targeting learning we have used:

   - a clicker with two buttons with different tones, which allowed to work simultaneously with two mice;
   - a target – a metal rod 33cm long ended with a yellow ball 2cm in diameter;
   - a plastic spoon (feeder), on which the rewards were given;
   - rewards – pieces of shelled sunflower, walnuts and seeds of hemp and millet.

The experiment was conducted according to the scheme below:

Phase I - familiarization with getting the food from the feeder.

On the beginning the feeder was moved in a way that mice could notice it and reach for the rewards. Then every feeder movement was preceded by the click: a higher tone for one mouse and a lower tone for the other. The procedure was repeated in two 3-minute sessions, after which the mice learned to recognize given sound as a reward announcement.

Phase II - targeting conditioning.

In this stage the target was introduced. If a mouse approached and touched it with its nose, the observer rewarded it. The conditioning of this behaviour was conducted in three 5-minute sessions.

Phase III - changing of the rewarding method.

After the targeting conditioning, 3 five-minute sessions took place during which the mice still received the positive reinforcement every time they have touched the target, however in each group the rewarding scheme was changed. The mice in the A group, which were put in the glass aquarium, still received the reward after each click, however the mice in B group (the plastic aquarium) received the reward only after the second or the third click.

Phase IV – extinction of the learned behaviour.

After the end of each session the reinforcement with the clicker and the rewards stopped. Next, the observer counted and noted the attempts that the mice took in order to acquire the reward. If the learned behaviour was not appearing after 3 minutes from the last attempt, it was considered extinct.

All stages were conducted by the same observer (N.S.). After finishing the experiment, both mice were marked with a marker and put back into the group aquarium from which another two mice were taken. In order to eliminate the influence of the type of the test aquarium, the group from the second round was introduced with reversed rewarding scheme: the mouse form the plastic aquarium received the reward after each time it touched the target, and the mouse form the glass one, after every 1, 2 or 3 attempts. In order to examine the significance of differences between the groups, we used the univariate analysis of variance.

Results and Discussion

The mice from the A group (which were rewarded after each touch of the target) attempted to acquire the reinforcement 6-29 times, whereas the mice from the B group, rewarded less frequently, tried to acquire the reinforcement 14-55 times. The average for the A group was: 15.6 ± 5.65 (SD), and for the B group: 31.6 ± 11.8. The difference between the groups was statistically significant (F=28.431, p<0.0001). The comparison of the attempts during the behaviour extinction is presented on Figure 1. During the extinction in some mice we noticed an intensification of the learned behaviour, e. g. grabbing or biting the target.

As it was mentioned in the introduction, there are two reinforcement schemes used in the clicker training of animals. In order to achieve the best possible training results, both schemes have to be used in the proper time (Sangha *et al.* 2002). If an animal is just starting to learn the given behaviour, the regular reinforcement is in order, so that the animal can quickly
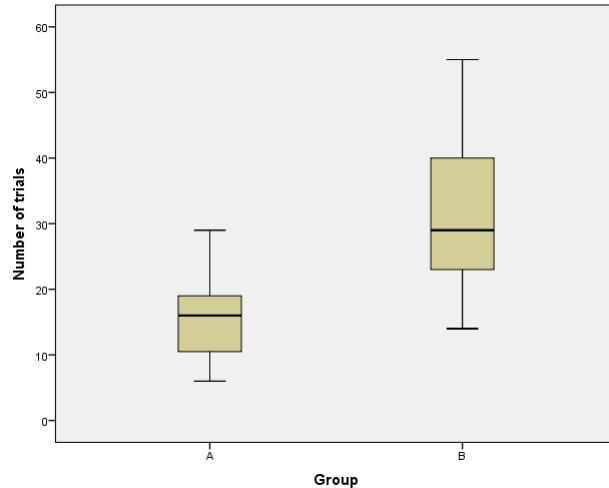
Figure 1. The number of attempts to acquire the reinforce-ment taken by the mice from the groups A (regular rein-forcement) and B (irregular reinforcement).

associate the reward with the behaviour and not get discouraged. It has been noticed also during our experiment - only three 5-minute sessions were suffi-cient to learn the targeting for all the mice. Wertheim and Singer (1964) and Williams (1989) believed that rewarding too rarely in the initial stage of training can cause a longer training time. After the condition-ing of given behaviour the rewarding scheme can be changed. Skinner (1938) and Humphreys (1939) wrote that using the partial reinforcement makes the learned behaviour more resistant to extinction, which they called the partial reinforcement extinc-tion effect. This effect was confirmed in this paper and in research on rats (D'Andrea, 1969) and inver-tebrates (Crancher *et al.* 1972; Sangha *et al.* 2002). Amsel (1972) suspected that the cause of this effect is a disruption taking place during the training. In oth-er words, if during the regular rewarding the animal does not receive the reward, the learned behaviour will extinct quickly, and during the partial reinforce-ment, the lack of reinforcements makes the animal frustrated, which causes increase of its attention. However, in the experiment of Sangha *et al.* (2002), where the snails were reinforced negatively, the au-thors pointed out that this theory in their case was incorrect because the lack of negative reinforcement was more beneficial for these animals and did not cause the frustration. Besides, the partial reinforce-ment introduces the element of surprise to the train-ing (Bouton and Sunsay, 2001). The animal cannot predict at which time it will receive the reward, so it becomes more vigilant and focused. Both theories

lead to the conclusion that the partial reinforcement increases the association of the behaviour with the reward. On the other hand Pryor (2002) assumed that periodical lack of the reward was a motivation for the animal to perform the learned behaviour in hope of getting the reward after the next attempt. Our experi-ment on mice showed that after the change of the rein-forcement scheme for the partial one, the learned be-haviour was more permanent, resistant to extinction. Smith and Davies (2008) in his research on dogs showed that during the time of extinction of the learned behaviour, the dogs which were reinforced with a clicker (with-out the reward) were more resistant to the extinction than the dogs from the control group which were not getting any reinforcement. It is not surprising because the dogs had associated the clicker with an announce-ment of the reward so they had been performing the learned behaviour until they got discouraged by the lack of the reward. On the other hand, Williams *et al.* (2004) showed that horses during the extinction phase did not show any difference between the clicker group and non-clicker group. Both experiments were conducted basing on a different methodology, also different from ours, so it is difficult to compare them. To conclude, in our paper we compared the resistance of mice for the behaviour extinction depending on the rewarding scheme. We showed that mice which were reinforced partially were more resistant to the extinction of behaviour than the mice rewarded for every performing of learned behaviour. As animals living in natural environment obtain rewards more or less regularly (which can be connected e.g. with stable/unstable food availability) results of our study could help understand also learning in nature.

REFERENCES

Amsel A. (1972). Inhibition and mediation in classical, Pavlovian and instrumental condition-ing. Pages 275-299 in: Inhibition and Learning. Boakes R.A. and M. Halliday, eds. Academic Press, London.

Bouton M. and Sunsay C. (2001). Contextual control of appetitive conditioning: Influence of a contextual stimulus generated by a partial reinforcement procedure. *Q. J. Exp. Psychol.* **54B,** 109-125.

Crancher P., King M.G., Bennett A. and Montgomery, R.B. (1972). Conditioning of a free operant in *Octupus cyaneus* Gray. *J. Exp. Anal. Behav.* **17,** 359-362.

D'Andrea T. (1969). Extinction of a heterogeneous chain after several reinforcement schedules. *J. Exp. Anal. Behav.* **12,** 127-135.

Dukas R. (2009). Learning: mechanisms, ecology and evolution. Pages 7-26 in Cognitive ecology II. Dukas R. and J.M. Ratcliffe, eds. University of Chicago Press, Chicago

Dickmann J., Gonzalez-Uarquin F., Reichel S., Pichl D., Radyshik K., Baumgart J. and Baumgart N. (2022) Clicker training mice for improved compliance in the Catwalk test. Animals 12, 3545

Humphreys L.G. (1939). The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. Exp. Psych.* **25,** 141-158.

Johnson M. (2003). Getting started: Clicker training for birds. Sunshine Books

Ohnishi H., Yamazaki S., Mochizuki K., Nakamura N. and Yuki K. (2003). Application of implicit learning to measurement of psychological effect of transmission delay in audiovisual communication. *Tech. Rep. IEICE.* **103,** 7-10.

Ohnishi H., Mochizuki K. and Yuki K. (2004). Psychological effect of pseudo response in audio communication with delay: Application of verbal conditioning. *Tech. Rep. IEICE.* **104,** 1-6.

Pryor K. (2002). Don't Shoot the Dog!: The New Art of Teaching and Training. Omnia Books Ltd., Scotland.

Ried J.P.(1996). Excelerated learning. James and Kenneth Publishers, Berkeley, CA.

Sangha S., McComb C., Scheibenstock A., Johannes C. and Lukowiak K. (2002) The effects of continuous versus partial reinforcement schedules on associative learning, memory and extinction in *Lymnaea stagnalis*. *J. Exp. Biol.* **205,** 1171-1178

Skinner B.F. (1938) The behavior of organisms: an experimental analysis. D. Appleton-Century Company Inc., New York.

Smith S.M. and Davis E.S. (2008). Clicker increases resistance to extinction but does not decrease training time of a simple operant task in domestic dogs (*Canis familiaris*). *Appl. Anim. Behav. Sci.* **110,** 318-329

Wertheim G.A. and Singer R.D. (1964). Resistance to extinction in the goldfish following schedules of continuous and variable interval reinforcement. *J. Exp. Anal. Behav.* **7,** 357-360.

Williams B.A.(1989) Partial reinforcement effects on discrimination learning. *Anim. Learning Behav.* **17,** 418-432.

Williams J.L., Friend T.H., Nevill C.H. and Archer G. (2004). The efficacy of a secondary reinforcer (clicker) during acquisition and extinction of an operant task in horses. *Appl. Anim. Behav. Sci.* **88,** 331-341.

Yamamoto M., Kikusui T. and Ohta M. (2009). Influence of delayed timing of owners' actions on the behaviors of their dogs, *Canis familiaris*. *J. Vet. Behav.* **4,** 11-18