Role of Unconditioned Stimulus Prediction in the Operant Learning: a Neural Network Model

S.E. Lew[1], C. Wedemeyer[2], and B.S. Zanutto[1,3]
[1]Fac. Ing. Universidad de Buenos Aires, [2]INGEBI-CONICET, [3]IBYME-CONICET,
1428 Buenos Aires, Argentina. E-mail: silvano@fi.uba.ar

Abstract

A neural network model of operant conditioning for appetitive and aversive stimuli is proposed. From neurobiological and behavioral bases it is assumed that animals are able to compute the prediction of the unconditioned stimulus. The prediction controls the learning of the correct response to obtain reward and to avoid punishment. The model has as inputs: all the conditioned stimuli and the unconditioned stimulus. The outputs are all the possible responses of the animal; each one is computed by one neuron. Based on Hebbian or anti-Hebbian learning, depending on the prediction, the synaptic weights of the response neurons are calculated. The synaptic weights of the neuron computing the prediction are calculated based on the Rescorla-Wagner model. The simulated and experimental data have been compared, showing that the model predicts relevant features of operant conditioning. This model is a theory of operant conditioning and provides principles to design autonomous systems.

1 Introduction

Behavioral experiments suggest that learning is driven by changes in the expectation about the future salient events, mainly reward and punishment. In operant and classical conditioning, the conditioned stimulus (*CS*) anticipates the unconditioned stimulus (*US*). Rescorla and Wagner [1] proposed that animals learn comparing what they expect on a given situation and what actually happens. As Staddon [2] has pointed out, animals act as the *CS* allows them to elaborate an expectation or prediction of the unconditioned stimulus. Furthermore, there are neural substrates of prediction and reward, such as the involvement of dopamine neurons of the ventral tegmental area (VTA) and sustantia nigra, identified with the processing of prediction and reward [3].

There are many non-mathematical theories to explain operant conditioning; sometimes there is no-agreement on their hypotheses and about the role of the prediction, especially in theories to explain escape and avoidance. In the one-factor theory [4], the avoidance response is reinforced by the *US* (e.g. shock). Instead in the two-factor theory [5], the avoidance response is reinforced by the reduction in fear due to the lack of the fear-eliciting *CS*. Here, the fear could be interpreted as a consequence of the prediction. In the cognitive theory of avoidance [6] it is assumed that during the acquisition phase, animals develop expectancies depending on its responses. However, Seligman and his colleagues [7] have proposed that under certain circumstances, animals develop the expectation that their behavior has little effect on their environment, and that this expectation may generalize to a wide range of situations. They called this effect: learned helplessness. There are also non-quantitative theories to explain the appetitive data [8]. There are also a few mathematical theories to explain operant conditioning, but only a few of them are able to describe the most relevant experimental features for appetitive stimulus [9] and for the aversive stimulus [10][11]. A theory is presented in this paper to explain the experimental data of operant conditioning in both the appetitive and the aversive conditions from neurobiological and behavioral bases.

2 The neural network model

As was said before, the VTA is involved in the prediction of the *US*, here this effect is computed by one neuron. The VTA neurons are connected with the prefrontal cortex through mesocortical dopaminergic system [12]. In monkeys, the prefrontal cortex is a region for convergence of five corticocortical pathways originated in the primary somatic, auditory, visual, olfatory and gustatory areas. These pathways are relatively independent one from the other, until they reach the prefrontal cortex, an associative area [13]. In the primate, the prefrontal cortex is the origin of a cascade of reciprocal connective links that flow down from it, to premotor cortex and from there, to primary motor cortex [14][15]. The cortical effect of the lateral interaction is simulated based on Kohonen [16] studies. He simulated this effect on neural networks when external afferents have different discharge frequencies for each neuron. He found a cluster of neurons discharging at their maximum frequency and the others with no responses. In this model it is assumed that, each response is generated by a cluster of neurons simulated by one neuron representing the effect of the cascade of motor links. The model (Fig. 1) has as inputs: all the conditioned stimuli (*CS*s), the

0-7803-7044-9/01/$10.00 ©2001 IEEE

unconditioned stimulus (*US*) and the outputs are all the possible responses of the animal (*Rs*).

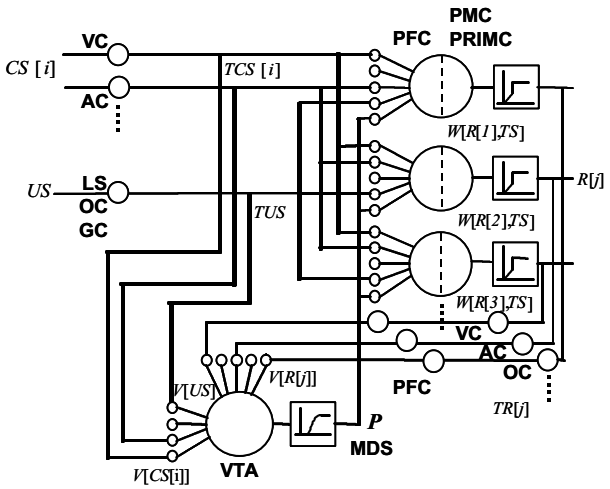


Figure 1: Neural network model, there is an artificial neuron computing the prediction (*P*), and one for each response *R[j]*. VC: visual cortex; AC: auditory cortex; OC: olfatory cortex; GC: gustatory cortex; LS: limbic system; PFC: prefrontal cortex; PMC: premotor cortex; PRIMC: primary motor cortex; MDS: mesocortical dopaminergic system; VTA: ventral tegmental area. The traces representing the short-term memory of the conditioned stimuli (*TCS*[*i*]), the unconditioned stimulus (*TUS*), and the responses (*TR*[*j*]) are inputs of the neurons computing *P*.and *R[j]*. The synaptic weights *V*[CS[*i*]], *V*[R[*j*]] and *V*[*US*] represent the associations between the inputs and *P*. The synaptic weights *W*[*R[j],TS*], are associations between *P*, *TCS*[*i*] and *TUS*, and the response neurons.

In this model all stimuli generates a short-term memory given by:

$$TS_n = TS_{n-1} \cdot (1-\alpha) + \alpha \cdot S_{n-1}$$
$$TR[i]_n = TR[i]_{n-1} \cdot (1-\beta) + \varepsilon \cdot (1 - TR[i]_{n-1}) \cdot R[i]_n$$

Where *S* is: *CS*[*1*], *CS*[*2*]…, *or US...; n* is the time unit. Here Greek letters represents constants. When S_n is greater than 0, $\alpha=\varepsilon$, and when $S_n=0$, $\alpha=\beta$.

The prediction (*P*) is calculated as:

$$P_n = \xi \Big/ (1 + e^{-\upsilon(X_n - \sigma)})$$
$$X_n = V[US]_n \cdot TUS_n + \sum_{i=1}^{NCS} V[i]_n \cdot TCS[i]_n + \sum_{j=1}^{NR} V[j+NCS]_n \cdot TR[j]_n$$

Where *V* represents the synaptic weights of the neuron that computes the prediction, these are the associations between the inputs and *P*.

i: conditioned stimulus index and j: responses index
NCS: is the number of conditioned stimuli,
NR: is the number of the possible responses.

The inputs to response neurons are: *P*, *TCS*s and *TUS*. The output of the response neurons (*R*[*j*]) is calculated as:

$$R[j]_n = f_2(Y[j]_n)$$
$$Y[j]_n = W[j][P]_n \cdot P_n + W[j][US]_n \cdot TUS_n + \sum_{i=1}^{NCS} W[j][i]_n \cdot TCS_n[i] + noise(n)$$

Where W represents the synaptic weights of the neuron that computes *R*[*j*], and *noise(n)*: white noise (amplitude=1/32). f_2 is a sigmoid given by:

$$f_2 = 0 \; if \; Y[j]_n < 0;\; f_2 = 1 \; if \; Y[j]_n > \mu;\; else \; f_2 = Y[j]_n$$

i: conditioned stimulus index
μ : is the threshold.

The animal executes the response *R[j]* when *Y[j]* is greater than μ.

If an animal is trained to respond with an appetitive amount of reward, when the amount diminishes, the expectation is modified in such a way that if the reinforcer is still appetitive, animals can stop responding [17]. This effect (some times called Crespi effect) is simulated modifying the Rescorla-Wagner model, considering not only the reinforcer value, but also the changes in its value. This was simulated adding to the reinforcer value the difference between what the animal received in the short term and in the medium term, modulated by a sigmoid. The modulation is to prevent the difference from affecting the intra-trial learning, since the Crespi effect is an inter-trial effect. This means that, if the animal receives a reinforcer of lower value, the summation of both terms is lower than the value of the actual reinforcer. The medium (TUS_{med}) and the short-term traces (TUS_{shor}) of the *US* are given by:

$$TUS_{med\,n} = TUS_{med\,n-1} \cdot (1-\delta) + \delta \, US_n$$
$$TUS_{short\,n} = TUS_{short\,n-1} \cdot (1-\rho) + \rho \cdot US_n$$

$$DUS = TUS_{shortn} - TUS_{medn}$$

The synaptic weights (*V*) of the prediction neuron (bounded between –1 and 1) are calculated as follows:

$$V[S]_n = (2 \Big/ (1 + e^{-\kappa \cdot VX[S]_n})) - 1;$$
$$VX[S]_n = VX[S]_{n-1} + \eta \cdot TS_n \cdot (US_n + \chi \cdot (\tau + DUS) \cdot (tanh(\gamma \cdot DUS))^{10} - X_n)$$

Where η can take two values, for the appetitive case, η_i controls the rise and η_d controls the decay of $VX[S]_n$. Animals learn faster when the stimulus is aversive than when it is appetitive. To simulate this effect, the constants to control the rise and the decay of VX are ζ times faster for the aversive case. Values are bounded between –10 and 10. The synaptic weight $V[US]$ of the prediction neuron is fixed to 0.1. From the two-factor theory point of view, this means that the *US* provokes fear, for the appetitive stimulus, it means, the animal raises its expectation for food. (Previous models included this hypothesis).

The synaptic weights of response neurons *j* are calculated by Hebbian or anti-Hebbian learning [18] depending on *P* as follows:

$$W[j][q]_n = \psi \cdot W[j][q]_{n-1} + \phi \cdot TQ_n \cdot TR_n[j] \cdot \Omega$$

Where *TQ* is: *P, TUS or TCS*[*i*] and the respective index q is: *P, US or i*; the first term includes a momentum [16]. For appetitive stimulus, if $P < \lambda$ then $\Omega = -\lambda$, otherwise $\Omega = \lambda$, the opposite for aversive stimulus. As was said above, we simulate the effect of learning that is faster for aversive than for the appetitive stimulus. To get responses other than escape and avoidance be fast dissociated, when *TQ*=*P*, ϕ has a value ϕ_{pav} higher than the value ϕ_a for all other cases. Also for appetitive stimulus, the association between food and the eating responses is one, to simulate that the animal eats if food is available.

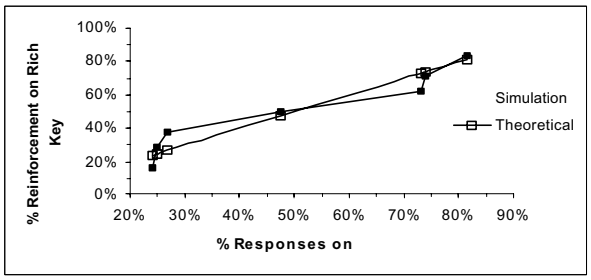
To simulate the animals exploratory behavior at the beginning of the experiment, the probability of generating random responses (*Pb*) decreases exponentially from a starting value (φ).

$$Pb_n = Pb_{n-1} \cdot \omega$$

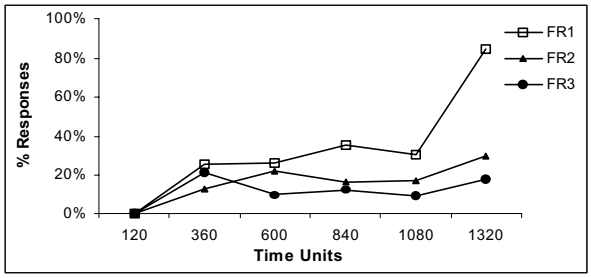
To simulate animals are starving at the beginning of a new appetitive conditioning experiment, φ has a value φ_{ap} higher than the value φ_{av} for the aversive condition. To simulate the animal increment of the exploratory behavior when a shock is applied, a response is chosen by random.

3 Computer Simulation

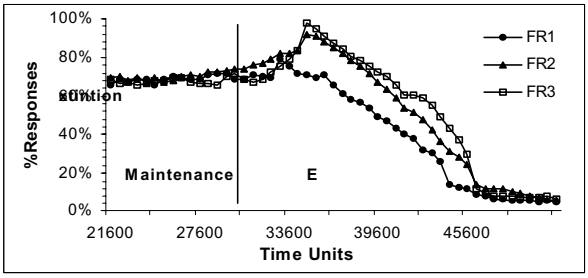
The simulated and experimental data for appetitive and aversive stimuli have been compared. In this model animals can make different responses: escape, avoidance, eating and others (grooming, vocalizations, etc.).



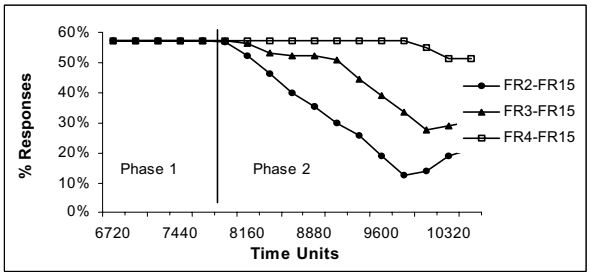
A



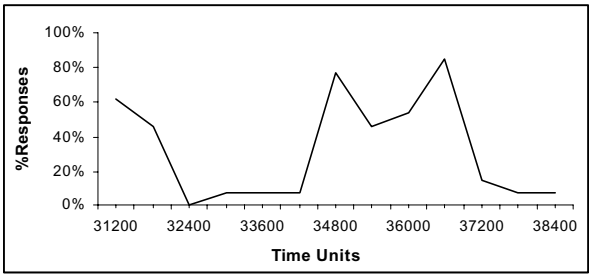
B



C



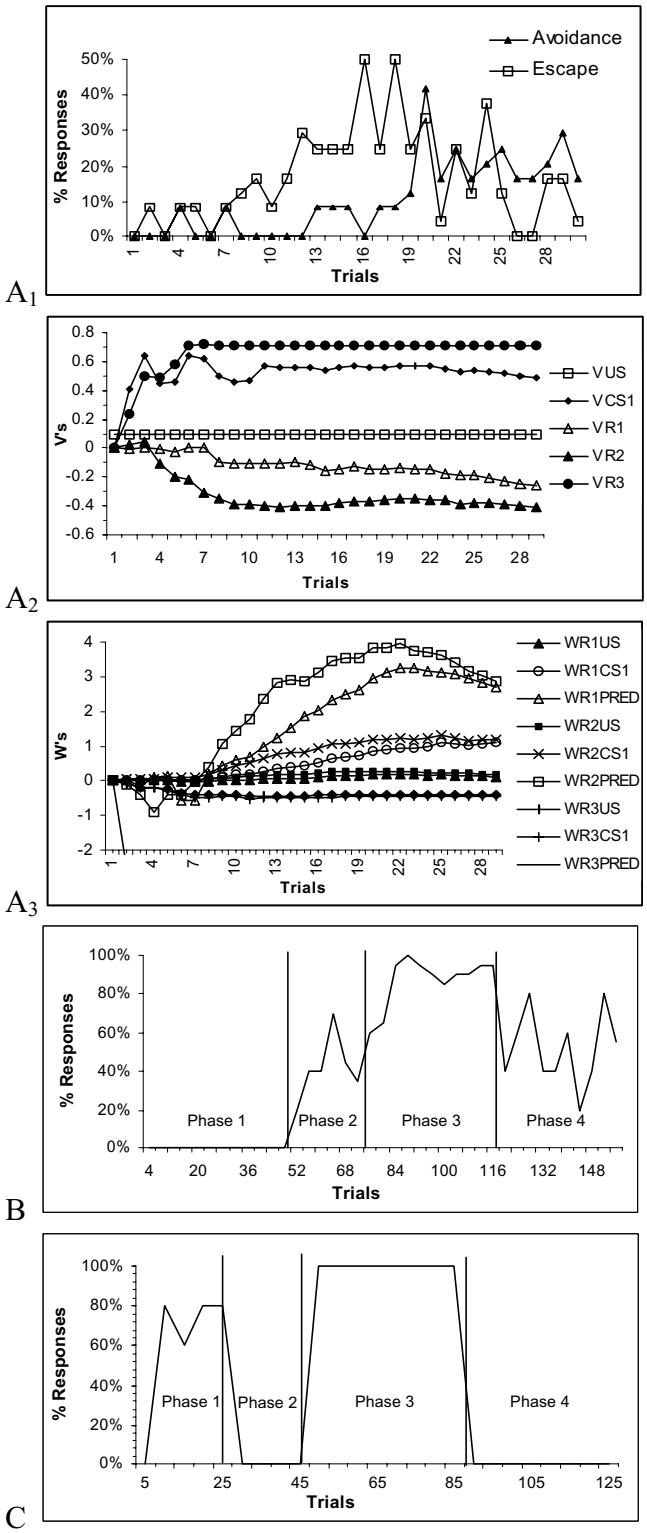
D

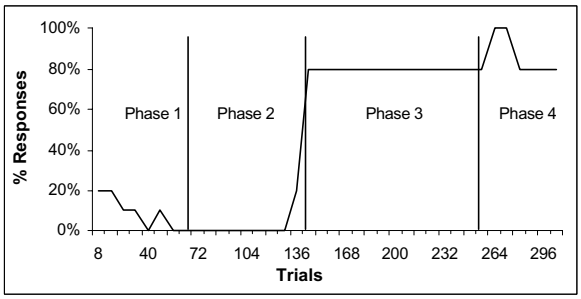


E

Figure 2 A: Matching law. The percentage of reinforcement on the rich key versus the percentage of responses on the same key is shown. Seven VI-VI schedules are simulated with the following probabilities: 1-0.2, 0.5-0.2, 0.33-0.2, 0.2-0.2, 0.2-0.33, 0.2-0.5, 0.2-1. B: Response selection. Three different conditions are simulated, one lever has a fix probability (VR 100) and the other has variable probability (VR 1, VR 2 VR 3). The figure shows that responses reinforced at a higher rate are selected faster. In this simulation *US*=6 to show clearer the effect, but it is present with the same value of *US* of the other experiments. C: Partial reinforcement extinction effect. A smaller percentage of reward (FR 3) in acquisition leads to a greater resistance to extinction. D: Successive negative contrast effect. Two different situations are shown: a) a small reinforcement (FR 15) for the two phases (control condition) and b) a large reinforcement in phase 1 (FR 2, FR3 and FR 4) is followed by small reinforcement (FR 15) in phase 2. In the second phase, for greater reinforcement the lower is the percentage of responses. E: Spontaneous recovery. Number of responses made in a session of extension as a function of the interval between the end of the acquisition session and the beginning of the extinction.

For the appetitive stimulus the model simulates the animal behavior in an experimental chamber, the operant response to get reinforcement is to press a lever and to reach the goal in a runway. In the simulation the values of *CS* =1 and *US*=20. One of the general principles was proposed by Herrnstein [19] known as the matching law. It states that if two concurrent variable-interval schedules (VI-VI) are offered to an animal, the proportion of responses for one alternative should equal the proportion of responses delivered by that alternative (Fig 2A). Animals show response selection, during the acquisition of operant conditioning (e.g. VR-VR), animals trained with high probability ratios learn faster than those trained with low probabilities [20] (Fig. 2B). Also partial reinforcement extinction effect [21] is simulated. After animals in a partial reinforcement (PRF) schedule (i.e. the response is reinforced with some probability) are shifted to extinction, a smaller amount of reward in acquisition leads to a greater resistance to extinction (Fig. 2C). When subjects are shifted from a continuous-reinforcement (CRF) schedule (e.g. every response reinforced) to a PRF schedule, PRF response rate is depressed relative to a control group receiving reinforcement with the same PRF probability throughout the experiment (Fig. 2D). This effect is called successive negative contrast effect [17]. After the extinction procedure in operant conditioning, if a subject is returned to the experimental chamber at some later time, spontaneous recovery of the operant response will be observed (Fig. 2E) [22]. For the aversive stimulus the model simulates the behavior of animals in a shuttle box, each trial lasting 60 sec. An animal starts each trial in one side of the box and a *CS* is turned on, if after 10 sec. the animal has not crossed to the other side, a shock (*US*) is delivered until the animal crosses and the trial is finished, or until the trial is completed. In all the simulations where the escape response is trained, the *US* intensity value is 2 (this is the maximum).





D

Figure 3 A: Shows the simulation second by second of the learning of different escape and avoidance responses. *CS*[*1*] is the conditioned stimulus, *R*[*1*] is the avoidance response and *R*[*2*] is the escape response After learning the escape response *V*[*R*[*1*]] is the only negative association. The higher *W* are those associated with the escape response *R*[*1*]. Under these conditions, when the input signal is the *US*, the animal will escape. The percentage of correct responses and the synaptic weights *V* and *W* are shown. Each mark represents the mean value of 3 trials, the constants are: β=0.01, δ=.0001, ε=0.25, ϕ_a=0.002, ϕ_{pav}=0.07, γ=5, η_i=0.2, η_d=0.01,κ=0.02, λ=0.6, μ=0.35, ρ=0.0003, σ=0.4, τ=0.45, υ=10, ω=0.9993, ξ= 1.2, ψ=.993, χ=20, φ_{ap}= 0.99, φ_{av}=0.2, ζ=10. Constants are the same for all the simulations. **B**: Modulation of the avoidance response. In the simulation the animal is exposed to a *CS*[*1*] in the presence of a *US*, and to *CS*[*1*] and *CS*[*2*] in the absence of the *US*, while the avoidance response (*R*[*1*]) is blocked. In the second phase the animal is trained in delay avoidance with *CS*[*3*]. Finally, the animal is tested with the *CS*[*3*] together with the excitatory *CS*[*1*] (phase 3), and *CS*[*3*] together with the inhibitory *CS*[*2*] (phase 4), but the *US* is never presented. In the first case the animal the percentage of responses is higher than in the second. **C**: Transfer of control from *CS*[*1*] to *CS*[*2*]. An animal trained to generate an avoidance response with *CS*[*1*]. In the following phase, while avoidance response is blocked, the unreinforced *CS*[*1*] and a reinforced *CS*[*2*] are presented. Finally, in the test phase, the animal is allowed to cross to the other side, whereas the *CS*[*1*] elicits avoidance responses (phase 3), *CS*[*2*] does not (phase 4). **D**: Learned helplessness and reversal. This experiment has four phases: in the first phase, the animals are exposed to an inescapable shock *US*. In the second phase the animal can not learn the escape response (*R*[*1*]). In the following phase the animal is forced to escape (*R*[*1*]). During the test phase, the animal learns the escape response

Learning delay avoidance is simulated when the escape and the avoidance responses are different [23][24] (Fig. 3A). *CS*[*1*] is the conditioned stimulus, *R*[*1*] is the avoidance response and *R*[*2*] is the escape response. The *V*[*R*[*1*]] and *V*[*R*[*2*]] are less that zero, *V*[*R*[*1*]] being the most negative. From the weights *W* , when *CS*[*1*] is applied, *R*[*1*] will be executed avoiding the *US.* Many experiments were carried out to prove that the avoidance behavior is explained by the two factor theory. Two of the most relevant are those of Rescorla & LoLordo [25], and Overmier & Bull [26]. In the first experiment, after the animal learned excitatory or inhibitory classical *CS-US* associations, presenting the *CS* during the performance of avoidance it can modulate the ongoing rates of the avoidance response. Fig. 3B shows the simulation, an animal is exposed to a *CS*[*1*] in the presence of an *US*, and to a *CS*[*1*] and a *CS*[*2*] in the absence of the *US*, while the avoidance response is blocked. Once the synaptic weights are stabilized, the animal is trained in delay avoidance with *CS*[*3*], but this phase is interrupted before the percentage of avoidance response reaches the maximum value in order to get positive modulation. Finally, the animal is tested with the *CS*[*3*] together with the excitatory *CS*[*1*] (where *V*[*CS*[*1*]] is positive, predicting *US*)., and *CS*[*3*] together with the inhibitory *CS*[*2*] (where *V*[*CS*[*2*]] is negative predicting no *US*), but the *US* is never presented. In the first case the animal almost always crosses, and in the second case the percentage of responses is lower. The second experiment mentioned above is shown in Fig. 3C. An animal is trained to generate an avoidance response with *CS*[*1*]. In the following phase, while the avoidance response is blocked, the unreinforced *CS*[*1*] and a reinforced *CS*[*2*] are presented. Finally, in the test phase, the animal is allowed to cross. While *CS*[*1*] does not elicit avoidance responses (*V*[*CS*[*1*]] association is extinguished), *CS*[*2*] does (*V*[*CS*[*2*]] is positively associated). Learned helplessness can be reversed: if after being exposed to the inescapable shock animals are forced to experience the escape contingency, animals learn to escape from a shock [27]. Fig. 3D shows the simulation of this procedure: the animal is exposed to the inescapable shock and in the following trials, the animal is forced to escape. In this way, the synaptic weights *V* associated with the escape response *R*[*1*] reach negative values, and all the *W* associate to *R*[*1*] reached slightly positive values. In this way, animals learn the escape response during the test phase.

4 Discussion

The simulated and experimental data have been compared, showing that the model predicts relevant features of operant conditioning for appetitive stimulus as: matching law, response selection, partial reinforcement extinction effect, spontaneous recovery, and successive contrast effect. In the appetitive case, whenever the prediction is higher than a threshold, the synaptic weights of the active response neurons increase their values according to the Hebbian law, meaning that the animal response is predicting *US*. If the prediction is lower than a threshold, the synaptic weights decrease (anti-Hebbian learning). The opposite rule is used for the aversive case. The synaptic weights of the prediction neuron associated with no reward responses are negative or zero; in this way when these responses are performed the prediction will be lower than when the reinforced responses are performed

(predicting no *US*). In the same way when avoidance or escape responses are performed, the associations between the prediction and these responses are negative (predicting no *US*). If the prediction is interpreted as "fear", the avoidance responses become conditioned inhibitors of fear [28]. We showed how the model explains experiments to support the two-factor theory. The model also explains how the experiment of Herrnstein & Hineline [29] to support the one-factor theory does not need a variable like fear or prediction. It is easy to see that if the function to switch from the Hebbian to the anti-Hebbian rule is controlled by a function of *US*, instead of by the prediction, then escape and avoidance can be learned without prediction [11]. However, in this case it is not clear how to learn other experiments (for example: learn modulation of avoidance and transfer of control without the prediction or a similar variable). This means that the one factor theory can explain some experiments, but not others. The model also explains experiments to support the cognitive theory and how under certain circumstances, animals develop the expectation that their behavior has little effect on their environment (Learn helplessness). Finally, this model also explains imitation in the same terms as Schmajuk & Zanutto [10].

5 Conclusions

Based on the prediction of the US, this theory explains the relevant features of operant conditioning for appetitive and aversive stimuli. We conclude that the dynamic process is the following: for the appetitive stimulus, responses are associated with the input stimuli by the Hebbian rule if the prediction of the unconditioned stimulus is higher than a threshold, and by the anti-Hebbian rule if it is lower. For aversive stimulus the criteria is the opposite.

6 References

[1] R.A. Rescorla and A.R. Wagner, A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: A.H. Black and W.F. Prokasy (Eds.) *Classical conditioning II: Current research and theory* (New York: Appleton-Century-Crofts, 1972)

[2] J.E.R. Staddon *Adaptive behavior and learning*. (Cambridge, Cambridge Univ. Press, 1983).

[3] W. Schultz, P. Dayan and R. Montague, A neural substrate of prediction and reward, *Science* 275, 1997, 1593-1598.

[4] R.J. Herrnstein, Method and theory in the study of avoidance, *Psychological Review*, 76,1969, 49-69.

[5] O.H Mowrer, On the dual nature of learning, an interpretation of conditioning and problem solving, *Harvard Educational Review*, 17, 1947, 102-148.

[6] M.E.P. Seligman and J.C. Johnston, A cognitive theory of avoidance learning. In: F. J. McGuigan and D. B. Lumsdenm (Eds.) *Contemporary approaches to conditioning and learning* (Washington D.C.: Winston-Wiley, 1973).

[7] J.B. Overmier and M.E.P. Seligman, Effects of inescapable shock upon subsequent escape and avoidance responding, *Journal of Comparative and Physiological Psychology*, 63, 1967, 28-33.

[8] A. Amsel. *Frustration theory*. (Cambridge, Cambridge Univ. Press, 1992).

[9] V. Dragoi and J.E.R. Staddon. The dynamics of operant conditioning. *Psychological Review* 106, 1999, 20-61.

[10] N. Schmajuk and B.S. Zanutto, Escape, avoidance and imitation: A neural network approach, *Adaptive Behavior*, 6, 1997, 63-129.

[11] B. S. Zanutto, S. Lew. A neural network model of aversive behavior. In: *Proceedings of the IASTED International Conference on Neural Networks*. Ed: M.H. Hamza. IASTED/ACTA Press. (Anaheim, Calgary, 2000). pp. 118-123.

[12] C.J. Pycock, R.W. Kerwin and C.J. Carter. *Nature*, 286, 1980, 74-6.

[13] Fuster, J.M (1997):The anatomy of prefrontal cortex. In: *The Prefrontal Cortex* , edited by Lippincott-Raven, pp.25. Philadelphia and New York.

[14] Watanabe-Saguaguchi, K. Kubota, K., and Arikuni, T. (1991): Cytoarchitecture and intrafrontal conecctions of the frontal cortex of the brain of the Hamadryas baboom (*Papio hamadryas*) *J. Comp. Neurol* : 311, 1991, 108-133.

[15] Barbas, H. and Pandya, D.N. (1989): Architecture and intrinsic connections of medial motor areas in the rhesus monkey. *J. Comp. Neurol*.,336: 1993,211-228.

[16] T. Kohonen. *Self-orgfanization and associative memory*. Springer-Verlag. 1988.

[17] L.P. Crespi. Quantitative variation in incentive and performance in the whit rat. *Am. J. of Psych*, 5, 1942, 467-517.

[18] D.O. Hebb, *The Organization of Behavior, a neuropsychological theory* (New York: Wiley, 1949).

[19] R.J. Herrnstein. Relative and absolute strength of response as a function of frequency of reinforcement. *JEAB*, 4, 1961, 267-272.

[20] J.E.R. Staddon and Y. Zhang. On the assignment-of-credit problem in operant learning. In: M.L. Commons, S. Grossberg and J.E.R. Staddon (Eds.). *Neural network models of conditioning and action* (Hillsdale, NJ, Erlbaum, 1991).

[21] A. Kacelnik, J.R. Kreps and B. Ens. Foraging in a changing enviroment: An experiment with starlings. In: M.L. Commons, A. Kacelnik and S.L. Shettleworth (Eds.), *Quantitative analyses of behavior VI: Foraging*. (Hillsdale, NJ: Erlbaum, 1987) pp. 63-87.

[22] N.J. Mackintosh *The psychology of animal learning*. (San Diego, CA: Academic Press, 1974).

[23] R.C. Bolles, Avoidance and escape learning: simultaneous acquisition of different responses, *Journal of Comparative Psychology and Physiological*, 68, 1969, 355-358.

[24] L.J. Kamin, The gradient of delay of secondary reward in avoidance learning, *Journal of Comparative and Physiological Psychology*, 50, 1957, 445-449.

[25] R.A. Rescorla and V.M. LoLordo, Inhibition of avoidance behavior, *J. of Comp. and Physiol.l Psych.*, 59, 1965, 406-412.

[26] J.B. Overmier and J.A. Bull, On the independence of stimulus control of avoidance, *Journal of Experimental Psychology*, 79, 1969, 464-467.

[27] M.E.P. Seligman, R.A. Rosellini and M.J. Kozak, Learned helplessness in the rat: time course, immunization and reversibility, *Journal of comparative and physiological psychology*, 88, 1975, 542-547.

[28] J.A. Gray, *The psychology of fear and stress* (London: Weidenfeld and Nicholson, 1971)

[29] R.J. Herrnstein and D.H. Hineline, Negative reinforcement as shock-frequency reduction, *Journal of the Experimental Analysis of Behavior*, 9, 1966, 421-430.