Neurocomputing 74 (2011) 3531–3542

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

# A dual association model for the extinction of animal conditioning

Ashish Gupta [a], Lovekesh Vig [b,*], David C. Noelle [c]

[a] Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA
[b] School of Information Technology, Jawaharlal Nehru University, New Delhi, India
[c] School of Engineering, University of California, Merced, CA, USA

ARTICLE INFO

*Article history:*
Received 7 October 2010
Received in revised form
8 June 2011
Accepted 9 June 2011
Communicated by I. Bojak
Available online 11 August 2011

*Keywords:*
Conditioning
Cognitive modeling
Renewal
Reinstatement
Savings

ABSTRACT

Reversal of synaptic plasticity has been the prevalent theory for extinction of animal conditioning. Phenomena like faster reacquisition after extinction are explained via residual synaptic plasticity in the relevant neural circuits. However, this account cannot explain many recent behavioral findings. This includes phenomena like savings in extinction, reinstatement, spontaneous recovery and renewal. These phenomena point to the possibility that extinction is not a mere reversal of the associations formed during acquisition. It instead involves the superimposition of some separate decremental process that works to inhibit the previously learned responses. We have explored this dual-pathway account using a neurocomputational model of conditioning. In our model, associations related to acquisition and extinction are maintained side by side as a result of the interaction between general neural learning processes and the presence of lateral inhibition between neurons. The model captures most of the relevant behavioral phenomena that prompted the hypothesis of separate acquisition and extinction pathways. It also shows how seemingly complex behavior can emerge out of relatively simple underlying neural mechanisms.



## 1. Introduction

The relationship between the learning of an association and the unlearning of that same association is commonly thought to involve a singular representation of the strength of association, with that strength rising during learning and falling during unlearning. In animal conditioning, this view suggests that the extinction of a behavior involves reversing the synaptic modifications made during the initial acquisition of that behavior. During acquisition training, the association between the conditioned stimulus (CS) and the unconditioned stimulus (US) is encoded by changing the strength of the synaptic interconnections between certain neurons in the brain. During extinction training, the changes made to these connections are reversed, causing the animal to stop producing the conditioned response (CR) [1–3]. While this theory is simple and elegant, it is not consistent with a growing body of behavioral findings.

Evidence from numerous studies points to the possibility that extinction is not a mere reversal of the associations formed during acquisition [4–15]. Phenomena like savings, reinstatement, spontaneous recovery and renewal suggest that extinction training involves the superimposition of some separate decremental process that works to inhibit previously learned response, leaving most of the originally acquired CS–US association intact. The phenomenon of savings [10–13,16,17] involves the relatively small amount of reacquisition training needed to restore the response after extinction training. In reinstatement [6,18], the response is restored through the presentation of US, alone. In spontaneous recovery [7,10,14,19] response to the CS is restored simply by the passage of time, after extinction training. Renewal [4,5] is said to occur when a shift in environmental context away from that in which extinction training took place results in renewed responding.

Recognition of retained association knowledge, even after responding has been extinguished, has led to theories involving residual synaptic plasticity and sub-threshold responding [1–3]. These theories hold that extinction training does not completely reverse synaptic changes made during initial acquisition, but only reverses these changes enough to effectively inhibit responding. When presented with the CS after extinction, the neural system involved in producing a response continues to become somewhat active, but not sufficiently active to produce an actual response. Thus, only small changes in association strength are needed to return this system to a state in which responding to the CS is robust.

However theories based on residual synaptic plasticity cannot account for some important additional observations. In particular, there is evidence that, just as extinction does not remove associations built up during previous acquisition training; subsequent reacquisition training does not remove the inhibitory force built

* Corresponding author.
*E-mail addresses:* ashish2240@gmail.com (A. Gupta), lovekeshvigin@gmail.com (L. Vig), dnoelle@ucmerced.edu (D.C. Noelle).

0925-2312/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2011.06.022

up during previous extinction training. For example, animals continue to show spontaneous recovery—a phenomenon that only arises after extinction training—even if they experience a subsequent period of reacquisition that removes the behavioral impact of the previous extinction process [10]. Also, just as reacquisition after extinction is faster than initial acquisition, subsequent extinctions are also faster than the first extinction [10,16,17].

Other phenomena, including conditioned inhibition, counter conditioning and feature positive discrimination shed further light on the nature of the associative changes during extinction. In conditioned inhibition [6], a CS is reinforced when presented alone, while it is not reinforced when presented in combination with a second CS. Due to this training, the second CS acquires inhibitory associations. In counterconditioning [6], a CS that was initially paired with one US is later paired with a contradictory US. This causes the extinction of the response corresponding to the first US. Finally, in a feature positive discrimination experiment [15], a CS is not reinforced when presented alone, while it is reinforced when presented in combination with a second CS. After this training, the strongest responding is observed for the CS combination. The second CS alone generates a somewhat weaker responding, while the first CS alone generates negligible responding.

In this paper, we show that the fundamental principles of neural computation, embodied in the Leabra modeling framework [20], spontaneously capture these phenomena of extinction. In particular, we show how synaptic plasticity, bidirectional excitation between cortical regions, and lateral inhibition between cortical regions interact, allowing the effects of previous acquisition and extinction to be maintained side by side. Of particular importance are processes of lateral inhibition, which introduce competition between neurons involved in the encoding of stimuli. Our model encodes the acquisition and extinction using a separate pool of neurons that compete with each other via the lateral inhibitory mechanism. During acquisition training, synaptic strengths change so as to encourage the activation of acquisition neurons and discourage the activation of extinction neurons. Similarly, during extinction training, synaptic strengths change so as to encourage the activation of extinction neurons and discourage the activation of acquisition neurons. As the extinction training proceeds, one by one, the acquisition neurons start losing the inhibitory competition. As soon as an acquisition neuron loses the inhibitory competition, its activation level drops dramatically, causing the synaptic modification process to effectively cease for it. Hence, much of the associational knowledge embedded in the synapses of the acquisition neurons is retained even after extinction. Similarly, many of the changes in extinction neuron synapses wrought during extinction training are retained after reacquisition training. Through this retention of synaptic strengths, our model is able to capture many of the behavioral results described above.

The paper is organized as follows. In Section 3, we present a brief description of the Leabra modeling framework. The details of our model are described in the subsequent section. That is followed by a detailed description of the different behavioral results along with the results of our simulation experiments. We conclude the paper with a general discussion of some relevant issues.

## 2. Related models

Temporal Difference Reinforcement Learning (TDRL) is unable to reproduce data confirming faster reacquisition after extinction due to the fact that it unlearns acquisition-related state values during extinction. Redish et al. [21] proposed an insightful variant of the TDRL model to address this issue. Their model introduced a state-classification process that determines the subject's current state and creates new states and state spaces when observation statistics change. The model hypothesized that tonically low expected reward produces a splitting of the state representation, with the different states being evaluated independently. However, the model leaves open issues such as the biological implementation of how the lack of expected reward would be signaled, and how cues would be categorized into situations within the brain. The proposed TDRL model captures the phenomenon of cued renewal by generating a new state on receiving consistently low reward during extinction. The renewal occurs because the associations formed during acquisition are not forgotten, and new extinction associations are formed in a different state. The Leabra model proposed in this paper also captures renewal with context units biasing the response of the subject in different contexts. This allows the network to generate different representations of an association for different contexts. Additionally, our model also captures the acceleration in learning observed during repeated reacquisition and extinction trials (Section 5.1), a phenomenon that the model by Redish et al. fails to capture.

More recently, Gershman et al. [22] extended the idea of state splitting to renewal and latent inhibition and suggested a latent cause model that allows for inference of new latent causes as the model gathers observations. Courville et al. [23] present a novel Bayesian account of change based on Pearce's theory of surprise in animal conditioning [24]. Their model infers that surprise expedites learning because it signals change, which consequently increases the uncertainty in the subjects' current beliefs about the world. An important difference between our model and these prior models is that we have built our model as a network of Leabra neural units, where each unit actually simulates neurons in the brain. The models mentioned above manage to capture much of the data presented in this paper, but these models tend to be mathematically abstract in nature, often leaving it unclear how these are to be translated into neural models.

The idea that acquisition and extinction involve learning in separate pathways has been proposed as an explanation to behavioral data by behavioral scientists [10,12] and has also been used in some highly successful neural models. Grossberg et al. [25] assume that extinction of conditioning happens not because of the reversal of learning in the on-pathway, but due to an active process of counter conditioning in the off-pathway. Our model does not 'assume' that lateral inhibition is sufficient for the emergence of independent pathways, rather it is a central finding. In contrast use of separate pathways for learning and unlearning was an assumption in Grossberg's work, implemented via dedicated learning rules which ensured that separate sets of weights change during learning and unlearning.

Another recent neural network model developed by Grossberg et al. [26]. simulates the interactions between different brain regions in an effort to identify the roles played by the amygdala, lateral hypothalamus and orbitofrontal cortex during conditioning. A relevant conclusion of their experiments is that none of the simulated brain regions are specifically involved in acquisition of conditioned responses. The paper further suggests that habit learning involves brain structures that are relatively independent of the brain structures involved in valuation like the amygdala and orbitofrontal cortex. This is consistent with our model's prediction that the biologically grounded property of lateral inhibition is sufficient for the emergence of separate independent pathways for learning and unlearning.

Thus, it may be possible for conditioning-related phenomena to emerge even in the absence of any specific brain architectures. We have kept our network architecture generic i.e. a standard three-layer network with an input layer, a hidden layer and an output layer. This is not to suggest that specific brain areas do not participate in the emergence of conditioning behaviors in specific animals, however our model is consistent with findings that

suggest that similar behavior should be exhibited in even the most primitive of brains [27–30].

## 3. Background

Our proposed model utilizes the well-known Leabra cognitive modeling framework [20] for developing a connectionist model that is grounded in known neurobiological principles. This section highlights the relevant properties of the Leabra framework that enable the development of biologically plausible connectionist models. Subsequently, the model proposed for demonstrating extinction phenomenon is described in detail.

### 3.1. Leabra modeling framework

The Leabra framework offers a collection of integrated formalisms that are grounded in known properties of cortical circuits but are sufficiently abstract to support the simulation of behaviors arising from large neural systems. The framework has been used to model a broad range of cognitive processes, including aspects of perception, attention, language, learning, and memory. Leabra includes dendritic integration using a point-neuron approximation, a firing rate model of neural coding, bidirectional excitation between cortical regions, fast feedforward and feedback inhibition, and a mechanism for synaptic plasticity that incorporates both error-driven and Hebbian learning. Of particular relevance to our model is Leabra's lateral inhibition formalism.

The effects of inhibitory interneurons tend to be strong and fast in cortex. This allows inhibition to act in a regulatory role, mediating the positive feedback of bidirectional excitatory connections between brain regions. Simulation studies have shown that a combination of fast feedforward and feedback inhibition can produce a kind of "set-point dynamics", where the mean firing rate of cells in a given region remains relatively constant in the face of moderate changes to the mean strength of inputs. As inputs become stronger, they drive inhibitory interneurons as well as excitatory pyramidal cells, producing a dynamic balance between excitation and inhibition. Leabra implements this dynamic using a k-Winners-Take-All (*kWTA*) inhibition function that quickly modulates the amount of pooled inhibition presented to a layer of simulated cortical neural units, based on the layer's level of input activity. This results in a roughly constant number of units surpassing their firing threshold. The amount of lateral inhibition within a layer can be parameterized in a number of ways, with the most common being either the absolute number or the percentage of the units in the layer that are expected, on average, to surpass threshold. A layer of neural units with a small value of this $k$ parameter (e.g., 10–25% of the number of units in a layer) will produce sparse representations, with few units being active at once.

In our model, acquisition-related and extinction-related learning occurs in two distinct sets of neurons that compete with each other via this lateral inhibition mechanism. Indeed, it is lateral inhibition, in conjunction with Leabra's synaptic learning mechanism, that enables the retention of acquisition knowledge in the face of extinction training and the retention of extinction knowledge in the face of re-acquisition training. A more detailed account of the *kWTA* function used by Leabra is provided in the Appendix.

### 3.2. Biological plausibility

The individual neural units in the Leabra modeling framework mimic the basic electrophysiological properties of biological neurons. Leabra incorporates the electrostatic and diffusion properties of ions to compute membrane potentials in its simulated neural units. Leabra adheres to the biological constraint of separating excitatory and inhibitory inputs by only directly simulating excitatory neurons (i.e. pyramidal neurons of the cortex), while utilizing the *kWTA* function described in the previous section to simulate the inhibitory inputs produced by the inhibitory interneurons in the cortex. Thus, the *kWTA* function allows Leabra to simulate the biological constraint of separating excitatory and inhibitory inputs by only directly simulating excitatory neurons. It also allows for smoother, more consistent activation dynamics than models that use direct inhibition between units, and is more robust while sustaining distributed representations [20]. Leabra strikes a balance between computational tractability and biological plausibility by using a point neuron activation function [20]. The point neuron activation function adheres to known dynamics of information processing in real neurons and allows for dendritic integration, thereby allowing for a reduction in the spatial geometry and improved computational tractability.

Leabra models simulate the spiking of a large number of neurons via a rate code approximation to discrete spiking. The output of a neural unit is a real-valued number that simulates the instantaneous rate at which an otherwise equivalent spiking neuron would produce spikes. The rate code approximation is supported by the fact that cognition is not impaired greatly by disruptions in the detailed firing properties of biological neurons. Also there is little evidence to support the fact that individual neurons communicate via precise spike timing, as that would be contradictory to a learning algorithm that is robust to noise in the spike timing.

The bidirectional connectivity between the hidden and output layers is crucial to the learning properties of our model. The Leabra learning algorithm adopts two phases of activation (please refer to Appendix for details). The difference between the two phases of activation is a measure of the units' contribution to the overall error signal. Bidirectional connectivity avoids the problems with backpropagation of computing error information and sending this error information backwards across synapses. Both parts of the signal are naturally propagated throughout the network via bidirectional connectivity. The activation states are local to the synapse where the weight changes must occur, and the biological plausibility of these weight changes in brain synapses is grounded in the properties of the GeneRec learning rule used by Leabra. The learning rule allows an error signal occurring anywhere in the brain to be used to drive learning everywhere, thereby allowing error signals to originate from different sources. In addition this form of learning is compatible with and in fact dependent on the bidirectional connectivity known to exist throughout the cortex [31,32]. Another advantage of GeneRec is that this difference in activation of a hidden unit during plus and minus phases is a good approximation for the difference in the net input to a hidden unit during plus and minus phases, multiplied by the derivative of the activation function [20].

## 4. The model

The three-layer Leabra network shown in Fig. 1 was used for our simulations. This model is an extension of the traditional Rescorla–Wagner [33] model of animal conditioning. In our model, a CS was encoded as a single input unit. The stimulus was recoded over the firing rates of 40 units grouped into a hidden layer. This hidden layer incorporated strong lateral inhibition, using a *kWTA* parameter of $k=5$, encouraging only 5 of the 40 units to be active at any one time. The hidden layer had a bidirectional excitatory projection to the output layer. The output layer contained 7 units, with $k=5$. Thus, our model used *kWTA*-based lateral inhibition in both the hidden layer and the output layer.

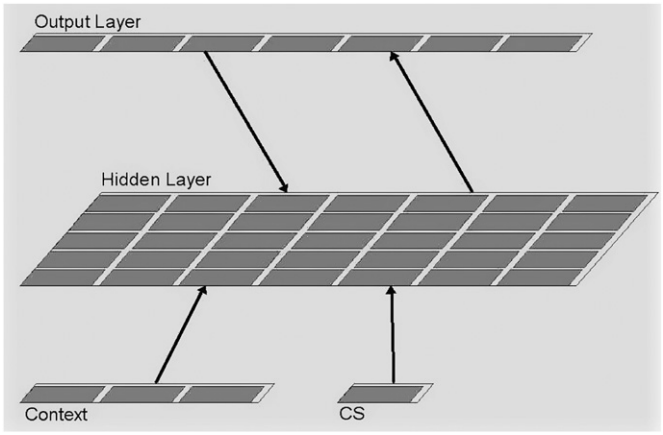


**Fig. 1.** The Leabra model used for our simulations. Each gray box corresponds to a neural processing unit. Each arrow represents interconnectivity between the units in two layers.

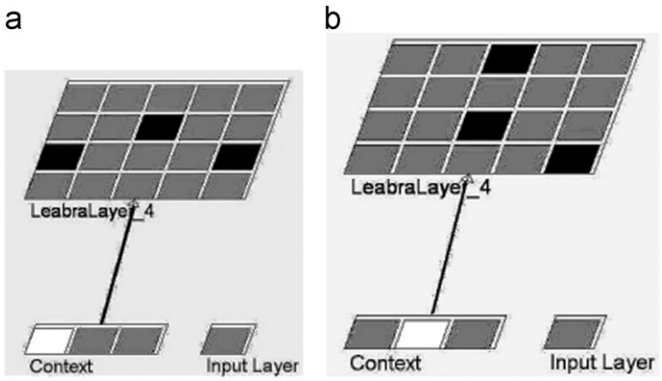


**Fig. 2.** Active hidden units in (a) context A and (b) context B.

For the simulation experiments that only used a positive reward, the first 5 units were interpreted as encouraging a positive response in the face of the stimulus, the average activation over these units determining the strength of the response. The final 2 units in the output layer encoded the "null response" generated after extinction training. These units offered a means to suppress the activity in the first 5 units via lateral inhibition. For fear conditioning, the last 2 units encoded the positive response, while the first 5 units encoded the animal's "freezing" response in the presence of the CS. We have used more neurons to encode CS+ rather than CS−(in the positive reward case as well as the fear conditioning case) based on evidence supporting higher neural activity during CS+ as compared to CS−[34]. However the number of units used to represent CS+/CS− in the output layer is not a critical model parameter and similar results were achieved with different output unit ratios as well.

In some of the behavioral experiments simulated by our model, conditioning and extinction trainings are conducted in different contexts. Contexts are different experimental chambers with distinct background lights, sounds or odors. In our model, context was encoded by weakly activating a single unit (setting to a value of 0.25, maximum possible value being 1.0) from a pool of contextual input units. These units were connected to the hidden layer units via a random pattern of connectivity, with a 90% probability of each connection being formed. These connections were randomly initialized at the beginning of each simulation trial and remained unchanged until the end of the trial. The weights of these connections were also randomly initialized.

Fig. 2 shows how switching on a unit of the context layer results in some activity in the hidden layer. This activation, although small, gives some units in the hidden layer a bias for activation. The units that are activated are the ones that are more strongly connected from the context layer as compared to the other hidden layer units. Thus, during a training trial, if the context layer is switched on, these units will have a greater chance of winning the kWTA competition, and thus, participating in the learning process. For different contexts, different units get a bias, thereby increasing the chances that different sets of units get involved in learning in different contexts. A switching of context in the simulation experiments was achieved by the switching of the active unit in the context layer. For the experiments in which the context was never switched, a single unit in the context layer remained active throughout the simulation.

It should be noted that the context layer and the CS layer connect to the same hidden layer and interact with the same hidden units with a high probability. The connection weights from the CS to the hidden layer were also randomly initialized. When simulating more than one CS, each was encoded over a separate input unit and a separate layer of 40 hidden units. All of the hidden layers participated in bidirectional excitatory connections with a single shared output layer, identical to the one previously described. Thus, different stimuli could not be represented using shared neural resources. This amounts to an assumption that the stimuli are all highly dissimilar, with each activating different neurons in the brain. This simplifying assumption is not a critical feature of this model.

Leabra's default parameters were used in these simulations, with only a few exceptions. To accommodate the relatively small size of this network, the range of initial random synaptic weights was reduced ([0.0; 0.1] rather than the default range of [0.25; 0.75]) and learning rate for synaptic modification was set to a smaller value (0.005, half of the default of 0.01). Also, individual neuron bias weights were removed.

Leabra allows users to specify scaling factors that determine the influence of the different projections into a receiving layer. This is achieved by setting the wt_scale.rel parameter, with a higher value enhancing the contribution and lower value weakening the contribution of a set of projections. In our model, the strength of the backward projection from the output layer to the hidden layer was weakened by setting the wt_scale.rel parameter to a value of 0.05 instead of the default value of 1. This decrease was required because with the default strength, the backward projections strongly activated all the hidden layers in the network, even those for which the corresponding input stimulus was not provided. Modifications of these kinds are common in smaller Leabra networks.

A randomly initialized network was used for each training trial. Each training session was terminated when the sum squared error (SSE) between the network's output and the expected output patterns fell below a criterion value of 1. All simulation experiments were repeated 25 times, and mean results across these runs are reported.

## 5. Experiments

The current section reports simulation results capturing behavioral data from phenomena like Savings, Reinstatement, Renewal, Feature Positive Discrimination and Conditioned Inhibition. A key feature of the Leabra framework that allows for these results is lateral inhibition. Lateral inhibition allows for sparse representations of network input–output mappings. This enables different sets of neurons to be active during acquisition and extinction and is critical to the savings experiments described in Section 5.1. Furthermore, the error-driven GeneRec learning algorithm used by Leabra allows the network weights to be updated in a biologically

plausible manner, thereby allowing the network to learn associations during acquisition and extinction. Finally, the role of context has been highlighted previously in Section 4 and the context units in our model play an important role in the Renewal, Reinstatement experiments as these bias different neural pathways to be active in different contexts.

### 5.1. Savings

It is generally seen that a relatively small amount of reacquisition training is needed to restore responding after extinction training. Rescorla [11,13] identified two different mechanisms that might be responsible for faster reacquisition of responding after extinction. First, it is possible that the association with the CS is not completely removed by extinction training—that residual synaptic plasticity retains some associational connection. Second, it might be the case that extinction training triggers faster subsequent learning—that a CS undergoing retraining is particularly quick to acquire new associative connections with the US due to its prior history. To investigate these two alternatives, Rescorla conducted the following experiments.

In one experiment, two stimuli, A and C, were initially trained and then extinguished. Two other stimuli, B and D, were presented without reinforcement. Once A and C were extinguished, A and B then each received the same number of conditioning trials, encouraging responding to these stimuli. At the end of this training sequence, A elicited stronger responding than B. This is a demonstration of savings, since A was previously acquired and extinguished and B was not. This observation does not distinguish between Rescorla's two alternatives, however. The A stimulus could have begun reacquisition training with some residual synaptic plasticity or the re-acquisition process could have operated at a faster rate for A. In order to separate these hypotheses, Rescorla tested responding to the compound stimuli AD and BC. Any residual synaptic plasticity in A should also be present in C, so responding to these two compounds should be roughly equivalent if both A and B grow equally in associational strength during reacquisition training. If, however, an association to A is learned faster because of its previous extinction, then greater responding should be seen to the AD compound. Surprisingly, neither of these outcomes was observed. Responding to BC was stronger than responding to AD. Rescorla concluded that A's dominance over B was the result of residual synaptic plasticity, and he explained the dominance of the BC compound in terms of a blocking-like effect. If associative change is governed by an error correction learning mechanism, and if stimulus A begins reacquisition training with a "head start" over stimulus B, there will be less error when stimulus A is presented, so the associational strength for A will grow more slowly than that for B. Since A's residual synaptic plasticity is shared by C, and since B's associational strength grows faster than that of A during re-acquisition training, the BC compound dominates over AD [11].

This explanation gave rise to a question: would A or B show greater associative change if the error signal during reacquisition training was equilibrated between them? In another experiment, Rescorla addressed this question by presenting the AB compound stimulus, rather than A and B separately, during reacquisition training [13]. When this was done, greater responding was generated by the AD compound than by the BC compound. Hence, Rescorla concluded that, in addition to leaving residual associative strength, extinction also causes the stimulus to gain new associative strength at a faster rate when it is, once again, reinforced. Through a similar set of experiments, he concluded that a stimulus that was previously extinguished and reacquired is more sensitive to subsequent non-reinforcement.

#### 5.1.1. Simulation 1A

Our first simulation experiment was designed to uncover the degree to which our model exhibits savings. Recall that animals are faster to reacquire an extinguished behavior, as compared to initial acquisition, and they are faster to extinguish a reacquired behavior, as compared to initial extinction. A randomly initialized network was trained to respond upon the presentation of the CS (A+). Once this training reached criterion, the network was trained to not-respond upon the presentation of the CS (A−). This pattern was repeated five times. Fig. 3 shows the number of trials required for successive acquisition and extinction trainings. Note that the required time quickly decreases. The model predicts that the required number of trials will asymptote to a small value after just a few acquisition-extinction iterations.

Why does this model exhibit savings? The network starts with small initial synaptic weights. Hence, a large change in weights is required for success during the first acquisition training session. During the first extinction training session, the weights to the acquisition neurons start decreasing and the weights to the extinction neurons start increasing. As soon as the extinction neurons win the inhibitory competition, the acquisition neurons tend to fall below their firing threshold. At this stage, the weights to the acquisition neurons stop decreasing, as these neurons are no longer contributing to erroneous outputs. Hence, a significant amount of acquisition-related association strength is retained through the extinction process. During the reacquisition training, the weights to the acquisition neurons increase once again and the weights to the extinction neurons decrease. Once again, the weights stop changing as soon as the extinction neurons lose the inhibitory competition. Hence, most of extinction-related plasticity is retained through the acquisition process. In this manner, subsequent acquisition and extinction trainings require a very small change in weights (Fig. 4). Effectively, acquisition and extinction associations are maintained side by side in the network, allowing for the rapid switching between them based on recent conditioning feedback.

#### 5.1.2. Simulation 1B

The design of this simulation experiment is shown in Table 1. As previously discussed, Rescorla designed this experiment to assess whether the rapidity of reacquisition was a result of residual synaptic plasticity or of an increase in acquisition speed after extinction [11]. A randomly initialized network was first trained on two CSs (A+ and C+) while two other stimuli were non-reinforced (B− and D−). Once the network reached criterion, it was then trained to extinguish A and C (A− and C−). During this session, B and D were presented in a non-reinforced manner as well (B− and D−). This was followed by training on A and B

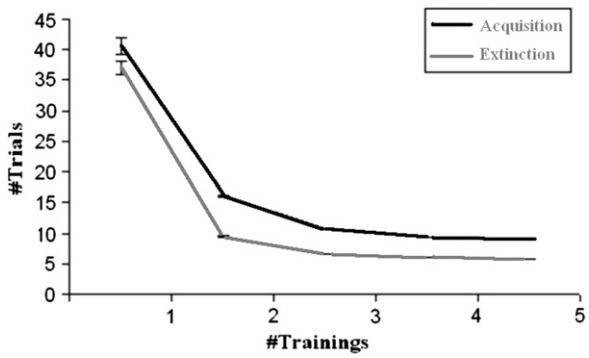


**Fig. 3.** Simulation 1A. The number of training trials required to reach criterion (*Y*-axis) decreases as the number of prior acquisition and extinction training sessions (*X*-axis) increases. Error bars report standard errors of the mean.

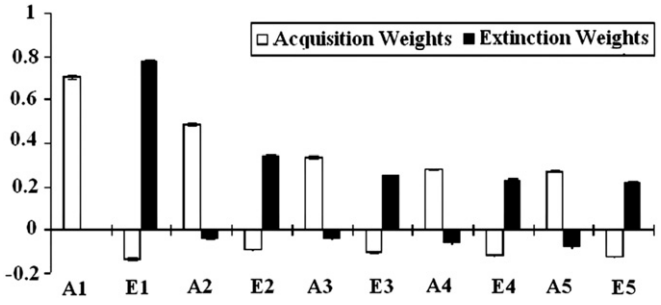


**Fig. 4.** This graphs plot the change in the summed connection weights in the acquisition pathway and in the extinction pathway (*Y*-axis) during the successive acquisition and extinction trainings (*X*-axis). The change in weights decreases in both the pathways as the number of prior acquisitions and extinctions training sessions increases. There seems to be a slow upward going trend in the weights in both the pathways, which appear to be a quirk of the simulator.

**Table 1**
The three training sessions, and single testing session, used in simulation 1B. Letters correspond to different stimuli. A plus indicates acquisition training, and a minus indicates extinction training.

| Acquisition | Extinction | Reacquisition | Test |
|---|---|---|---|
| A+ | A− | | |
| B− | B− | A+ | AD |
| C+ | C− | B+ | BC |
| D− | D− | | |

(A+ and B+) for 17 trials.[1] At the end of these training sessions, the response to A was much stronger than the response to B ($t(48)=2.517$, $p<0.0187$), as shown in Fig. 5. This is in accordance with Rescorla's observations. Finally, the network was tested on the compounds: AD and BC. As observed behaviorally, the network showed greater responding for BC than for AD ($t(48)=55.72$, $p<0.0001$) (see Fig. 6).

While the simulation results show a qualitative match to the behavioral data, they are more exaggerated. This pattern can be seen in many of the subsequent experiments as well. We report the activations as the sum of the output layer unit activation values. While Leabra units can take any real numbered value, in practice, it is generally seen that due to recurrent activation patterns, the default network parameters (including the parameters of the sigmoidal activation function) typically result in individual units either strongly responding or very weakly responding.

In addition to capturing these general results, our model also matches more subtle nuances in Rescorla's data. First, the compound BC produced a much stronger response in animals than either B or C alone. Second, the compound AD was found to produce significantly weaker responding than stimulus A, alone. Our model captures both of these results. How can these results be explained?

The compounds AD and BC are logically equivalent at the end of the extinction training. During the retraining phase, stimulus A quickly reacquired the responding due to its prior history of acquisition and extinction. The acquisition pathway for stimulus B also strengthened during this phase. However, the 17 training trials of this phase prove to be insufficient to strengthen the pathway enough to generate any noticeable gain in B's responding. At this stage, stimulus C had strong weights in both the pathways. However, due to the more recent extinction training, the extinction neurons won the inhibitory competition when C was tested alone. When C was presented in combination with B, B's additional support proved to be sufficient to push the acquisition neurons ahead in the inhibitory competition, causing a dramatic increase in responding. In comparison, the weakly extinction-biased weights of D, when combined with the strong but balanced weights of A, were enough to start to tip the inhibitory competition in the direction of a null response when A and D were combined.

#### 5.1.3. Simulation 1C

Table 2 shows the design of this simulation. As previously discussed, Rescorla designed this experiment to identify accelerated learning after extinction by equalizing for the amount of error experienced by both A and B during reacquisition training. A randomly initialized network was trained as before, with the only difference being the use of a compound stimulus (AB+) during reacquisition training. Reaching criterion during reacquisition required 17 trials, on average. As observed in animals, the network produced stronger responding for AD than for BC ($t(48)=12.225$, $p<0.0001$). See Fig. 7.

Our model contains no mechanism for increasing the rate of learning for A after its extinction. So, how did our model capture this pattern of performance? This was not due to a speeding of learning with regard to the A stimulus, but due to a blocking of learning with regard to the B stimulus. During the reacquisition phase, due to stimulus A's prior history of conditioning and extinction, the responding for the AB compound increased quickly. Hence, the error signal which was responsible for the weight update dropped quickly. Hence, the change in weights for stimulus B became smaller and smaller with each trial. The average weight increase for stimulus B during the retraining phase of this simulation was 0.212. In comparison, the average weight increase during the retraining phase of the previously described simulation was 0.342. This happens because in the last simulation, stimulus A and B were presented individually. Stimulus B, which had no prior conditioning experience continued to generate a very small response and hence received a large error signal throughout the training.

### 5.2. Reinstatement

The phenomenon of reinstatement involves the restoration of an extinguished response through the presentation of the US alone. Reinstatement has generally been reported in the domain of fear conditioning. Fear memory is believed to be stored in the amygdala [35] . In a typical experiment [18], the animals are first given baseline training, where they are trained for lever pressing (by rewarding them with food pellets). This is followed by fear conditioning, where the animals learn to suppress the lever pressing behavior in the presence of a CS, by pairing that CS with a foot shock. Then, the fear conditioning is extinguished via the omission of the foot shock. At the end of the extinction training, the animals are subject to a non-contingent foot shock without the presentation of the CS. When tested in the presence of the CS, it is observed that the lever pressing behavior is suppressed[2] once again.

Many theories have been proposed to explain this phenomenon. One theory states that during the acquisition training, along with the CS, the US is also encoded as part of the stimulus [6,18]. Hence, when the US is presented alone, the strength of the response-related association would still change, resulting in reinstatement. Another theory states that the background

[1] This number of trials was chosen to make these results comparable to those from simulation 1C. In simulation 1C, it was found that 17 trials were needed, on average, to train the AB compound to criterion.

[2] Suppression is measured as a ratio; suppression ratio=(responding in the presence of CS)/(responding in the presence of CS+baseline responding).

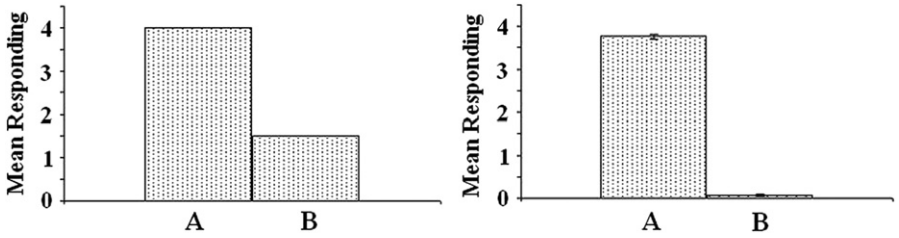


**Fig. 5.** Simulation 1B. Left: results reproduced from Rescorla's experiment—mean responding for A and B after the reacquisition phase. Right: simulation result—response magnitude for A and B at the end of the reacquisition phase, with error bars showing standard errors of the mean.

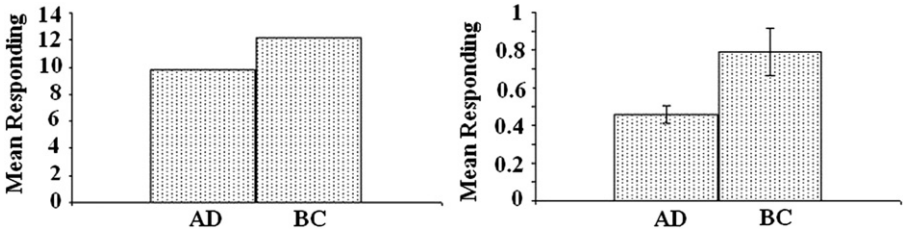


**Fig. 6.** Simulation 1B. Left: results reproduced from Rescorla's experiment—mean responding for AD and BC. Right: simulation result—response magnitude for AD and BC compounds, with error bars showing standard errors of the mean. Note that BC produced a stronger response than AD.

**Table 2**
The three training sessions, and single testing session, used in simulation 1C. Letters correspond to different stimuli. A plus indicates acquisition training, and a minus indicates extinction training. Note that "AB+" indicates that both A and B were presented together, as a compound, and this compound was reinforced.

| Acquisition | Extinction | Reacquisition | Test |
|---|---|---|---|
| A+ | A− | | |
| B− | B− | AB+ | AD |
| C+ | C− | | BC |
| D− | D− | | |

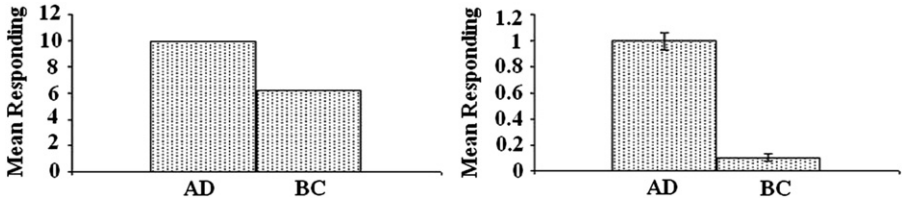


**Fig. 7.** Simulation 1C. Left: results reproduced from Rescorla's experiment—mean responding for AD and BC. Right: simulation results—response magnitude for AD and BC compounds, with error bars showing standard errors of mean. Note that AD produced a stronger response than BC.

context, in which acquisition and extinction trainings take place is encoded as a part of the CS [6]. Hence, the presentation of US alone would still change the response-related associations, resulting in reinstatement.

The theory that extinction forms a new memory and is not erased by conditioning is well supported in the literature. It has been shown that destruction of the ventral medial prefrontal cortex blocks recall of fear conditioning [36,37] hinting that the medial prefrontal cortex may be responsible for the storage of long-term extinction memory. Milad and Quirk [38] performed experiments to show that infralimbic neurons recorded during fear conditioning and extinction fire to the tone only when rats are recalling extinction on the following day. They further showed that stimulation resembling extinction-induced infralimbic tone responses is able to simulate extinction memory. Westbrook et al. [39] studied the role of context in reinstatement and their model suggested that reinstatement was specific to the context presented post-extinction.

*5.2.1. Simulation 2*

This experiment was performed to test if the non-contingent presentation of the US alone results in the reinstatement of an extinguished fear response. The design of this simulation experiment is shown in Table 3. A randomly initialized network was given baseline training, where it learned to press the lever (L−) stimulus. Then, the network was subjected to fear conditioning in the presence of a stimulus (LT+,L−). This was followed by the extinction of the fear conditioning (LT−, L−). Then, the network's responding was tested for in the presence of the CS (LT) and in the absence of the CS (L). This was followed by a single non-contingent presentation of foot shock (L+). Finally, the network's responding for L and LT combination was tested once again. We found that there was negligible change in the responding to L before and after the single shock stimulus presentation. The response magnitude before the shock presentation was $1.881(\pm 0.014)$ while after the shock presentation was $1.862(\pm 0.005)$ ($t(48) = 1.259$, $p < 0.2128$).

However, the responding for the LT combination showed a significant drop ($t(48) = 20.504$, $p < 0.0001$) (see Fig. 8).

Most theories of conditioning ignore the fact that the initial baseline training, where the animal learns to press the lever must entail the formation of some associations. In contrast, in our simulations, the lever stimulus (L) forms an integral part of the entire training process. During the baseline training (L−), the lever acquires strong food associations. These food associations survive through the fear conditioning, and L continues to elicit a strong food response. In contrast, after fear conditioning, the fear associations for the LT combination are stronger than the food associations, thereby eliciting a strong fear response. During the extinction training, the fear associations for the LT combination weaken only to the extent required to lose the inhibitory competition. Hence, a large proportion of these associations survive the extinction training. A single shock presentation (L+) results in a sufficient increase in L's associations to cause the fear associations for the LT combination to start winning the inhibitory competition once again. The food associations of L still remain strong enough to win the inhibitory competition when L is presented alone.

### 5.3. Renewal

In the phenomenon of renewal, a change of background context after extinction causes a robust return of conditioned responding. Bouton et al. [4–6] have performed a thorough investigation of several different variants of this phenomenon. In one of the versions, called the ABA renewal [5], the animals are first given baseline training (lever pressing) in two different contexts (context A and context B). This is followed by fear conditioning in context A. Then, the fear response is extinguished in context B. Finally, the fear response is tested in both the contexts. A greater fear response is observed in context A than in context B. As mentioned earlier, contexts are separate and counterbalanced apparatuses housed in different rooms in the laboratory that differed in their tactile, olfactory or visual aspects.

In another variant called the ABC renewal [5] the same training process as in ABA renewal is conducted. The only difference is that the final test of responding is conducted in an additional neutral context (context C), where the animal only received the baseline training. Animals display a greater suppression of the responding in context A as compared to context C. Moreover, they display a lesser suppression in context B than in context C.

Bouton has explored the possible explanations for these phenomena. One explanation is that the context forms a part of the CS and hence, forms associations with the US. However, Bouton et al. show that the renewal effect can occur even in the absence of demonstrable excitation in the contexts of training. Based on their experiments [4], they postulate that the context itself does not form any association. Instead, it modulates or "sets the occasion" for the formation of CS–US or CS–no US associations.

Wagner et al. [15] have proposed a real-time model that provides an interesting explanation for the role of context in renewal. In one of their experiments, a CS was presented for a long duration (just like the background context), followed by, and terminating with a short presentation of a second CS. This combination was reinforced, whereas, the second CS alone was presented without reinforcement. Through this training, the animals learned to respond to the sequential presentation of the two stimuli, but did not respond to either of the stimuli when presented alone. Their model is able to capture this phenomenon. In their model, each training trial is divided into smaller subtrials. The long duration presentation of the first CS is treated as multiple non-reinforced presentations of that stimulus. The final state of the training trial is treated as a reinforced presentation of the CS combination. With this training, the model easily fits the behavioral data.

#### 5.3.1. Simulation 3

This simulation demonstrates that shifting the animal to a context different from the context of extinction results in a renewal of the conditioned responding. The design of this simulation experiment is shown in Table 4. A randomly initialized network was given baseline training (L−) in three different contexts (A, B and C). After the baseline training, fear conditioning (L−, LT+) was conducted in context A followed by extinction (L−, LT−) in context B. Finally, responding with (LT) and without (L) CS was tested in all three contexts. In accordance with the behavioral results, the network exhibited a greater suppression in context A as compared to context C ($t(48) = 7.712$, $p < 0.0001$) (see Fig. 9). The suppression in context B was lesser than in context C ($t(48) = 2.21$, $p < 0.0368$) (see Fig. 10).

**Table 3**
The four training sessions used in simulation 2. L corresponds to the lever stimulus and T corresponds to a conditioned stimulus. A plus indicates fear conditioning, and a minus indicates extinction training. Note that "LT+" signifies that during the fear conditioning with T stimulus, the lever (L) stimulus was also present, and hence the LT compound was reinforced.

| Baseline | Conditioning | Extinction | Reinstatement |
|---|---|---|---|
| L− | L−<br>LT+ | L−<br>LT− | L+ |

**Table 4**
The three training sessions and a single testing session, used in simulation 3. A, B and C correspond to different contexts, L corresponds to the lever stimulus and T corresponds to a conditioned stimulus. + indicated fear conditioning and −indicates extinction of fear conditioning.

| Baseline | Conditioning | Extinction | Test |
|---|---|---|---|
| A: L−<br>B: L−<br>C: L− | A: L−<br>A: LT+ | B: L−<br>B: LT− | A: LT<br>B: LT<br>C: LT |

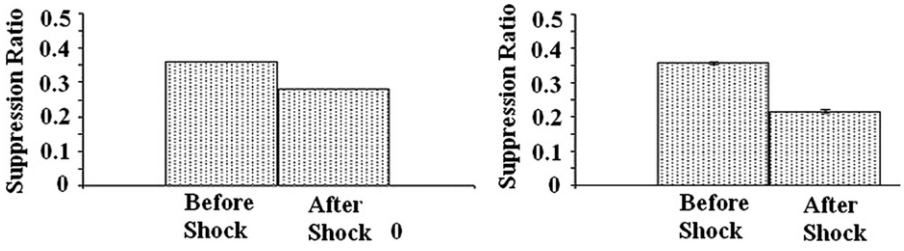


**Fig. 8.** Simulation 2. Left: results reproduced from Rescorla's experiment—suppression ratio in the presence of the CS, before and after reinstatement. Right: simulation results—suppression ratio in the presence of the CS before and after the presentation of the non-contingent shock. Note that there is a greater suppression of the response (smaller suppression ratio) after the shock presentation than before the shock presentation.

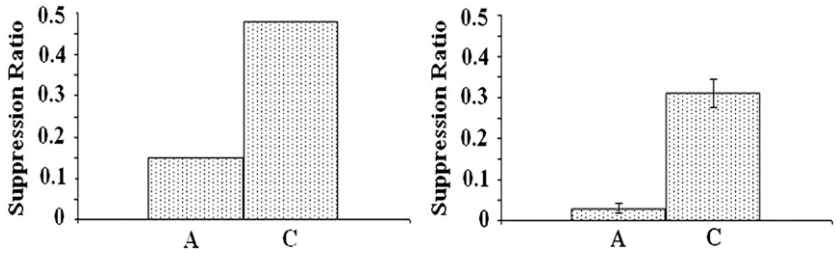


**Fig. 9.** Simulation 3. Left: results reproduced from Boutons experiment—suppression ratio in context A and context C. Right: simulation results—suppression ratio in context A and context C. Note that there is a greater suppression of the response in context A, the context of fear conditioning, as compared to context C, the neutral context.

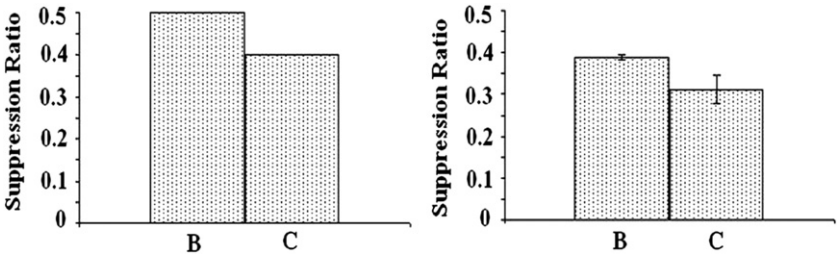


**Fig. 10.** Simulation 3. Left: results reproduced from Bouton's experiment—suppression ratio in context B and context C. Right: simulation results—suppression ratio in context B and context C. Note that there is a lesser suppression of the response in context B, the context where fear conditioning was extinguished, as compared to context C, the neutral context.

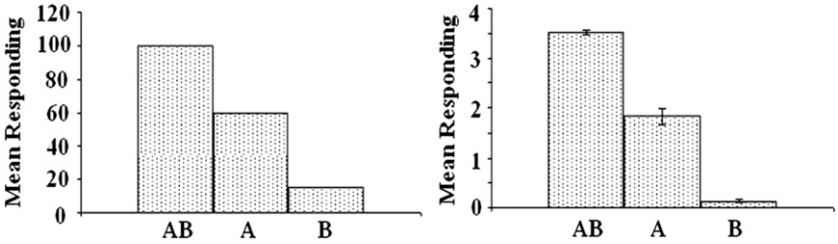


**Fig. 11.** Simulation 4. Left: results reproduced from Wagner's experiment—percentage responding for AB compound, A and B. Right: simulation results—response magnitude for AB, A and B. Note that the response magnitude is greater for AB compound than for A.

Fear conditioning is conducted in context A. Hence, A acquires fear associations. In contrast, due to extinction training, context B acquires extinction-related associations. Context C, the neutral context, does not acquire either associations. The CS undergoes acquisition training (in context A) and extinction training (in context B). Hence, both the acquisition-related and extinction-related associations strengthen for it. Hence, when the CS is combined with context A, the response tends to shift in favor of the fear response. When the CS is combined with context B, the response tends to shift in favor of extinction. With context C, the response remains intermediate. Additionally, the presentation of context alone fails to generate any activity in the output layer of our model. Hence, as observed by Bouton, the contexts of training do not acquire any demonstrable associations during this training process.

### 5.4. Feature positive discrimination

In feature positive discrimination [15], a CS is reinforced in combination with a second CS, and it is not reinforced when presented alone. With this training, the CS combination generates a very strong response. The responding to the first CS alone is negligible. The second CS alone exhibits an intermediate level of responding. The single association theory makes a contradictory prediction for this phenomenon. It predicts that the first CS, starting with a net zero association, would acquire a net association that is negative in magnitude. As a result, the association acquired by the second CS will have to be especially strong in order to overcome the subtracting effect of the first CS, when presented in combination. Hence, the responding should be stronger for the second CS alone as compared to the responding for the CS combination.

#### 5.4.1. Simulation 4

In this simulation experiment, a randomly initialized network was trained on a stimulus combination (AB+), while one of the stimuli was presented without reinforcement (B−). At the end of this training, the network's performance was measured for A, B and AB compound. In accordance with the behavioral data, the AB compound showed a stronger responding than A ($t(48) = 10.684$, $p < 0.0001$). (see Fig. 11).

Stimulus B, which participates in both acquisition and extinction trainings, strengthens its associations in both pathways. Hence, when B is presented alone, the two pathways cancel each other's effects via lateral inhibition, generating very small output. Stimulus A only strengthens its acquisition-related associations. As a result, when A is presented alone, it generates a strong response. AB combination generates an even stronger response due to the mutual support of acquisition associations of A and B.

### 5.5. Conditioned inhibition

In conditioned inhibition [6], a CS is reinforced when presented alone, but not reinforced when presented in combination with a second CS. Due to this training, the second CS acquires inhibitory capabilities—it can inhibit the responding, when it is combined with some other reinforced stimulus. The single association theory predicts this result. It posits that the second stimulus starting with net zero association would generate a net association that is negative in magnitude. Hence, when combined with the first stimulus, the summed strength of the association becomes smaller, resulting in a diminished responding.

#### 5.5.1. Simulation 5

The goal of this experiment was to test if our model predicts conditioned inhibition. A randomly initialized network was reinforced for two different stimuli (A+, B+), while stimulus A was presented without reinforcement in the presence of stimulus C (AC−). At the end of this training, the network was tested for the BC compound. As expected, the BC compound generated a weaker responding as compared to B. The mean responding for the BC compound was 1.459($\pm$0.184) while that for B was 4.367($\pm$0.045) ($t(48)=15.374$, $p<0.0001$). This happens because C acquires extinction-related association which inhibits B's responding.

## 6. Conclusions

We have proposed a neurocomputational model for the extinction of animal conditioning. Contrary to the notion that extinction involves only a reversal in previously acquired synaptic associations, the model posits the existence of a separate pathway that interacts with the acquisition-related pathway through the interaction of foundational neural processes, including error-driven synaptic plasticity, bidirectional excitation, and strong lateral inhibition. We have shown that our model captures the relevant patterns of performance exhibited by animals. Another strength of our model is that it does not depend on the specific properties of particular brain systems, such as the hippocampus or the cerebellum. Hence, it helps in explaining why vastly different brains produce similar patterns of learning.

What is the biological plausibility of the dual-association hypothesis? The inhibitory circuits that are responsible for reducing the expression of fear are not fully understood. However, there is evidence highlighting the role of the medial prefrontal cortex (mPFC) in memory circuits for fear extinction [35,40,41]. The ventral and dorsal mPFC have been shown to be differentially involved in acquisition and extinction of conditioned fear responses [36]. Several distinct nuclei of the amygdala have been shown to be differentially involved in acquisition, extinction and expression of fear responses [42–44]. Simulations of savings-related mechanisms in the cerebellum during eyelid conditioning reveal that extinction does not cause a reversal of plasticity in the acquisition-relevant brain areas [3]. This evidence when pieced together lends support to the notion that extinction is new learning, rather than erasure of conditioning. Moreover, the extinction pathway does not necessarily need to comprise of neurons solely dedicated to suppress the conditioned response. The neurons used for generating alternative responses could serve the dual purpose of simultaneously suppressing the conditioned response as shown in some of our simulations.

One unresolved issue is concerning the phenomenon of spontaneous recovery. Spontaneous recovery has traditionally been used as one of the strongest arguments in favor of the dual-association theory of conditioning. However, simply the use of two separate associations is insufficient to provide any explanation for the phenomenon. An understanding of how the passage of time influences the two associations would be required to explain spontaneous recovery. Spontaneous recovery seems to be selectively biased toward the recovery of the extinguished behavior. A closer examination of the phenomenon could thus shed more light on whether and how the learning that happens during acquisition and extinction differs.

Leabra's error-driven learning algorithm has been used for these simulations. The Leabra framework also implements the temporal difference learning algorithm [45]. This algorithm has been used to simulate several conditioning phenomenon where rewards are temporally delayed. Our future work will also focus on extending the model to simulate these type of phenomena as well.

## Acknowledgments

The authors would like to thank the two anonymous reviewers for their valuable feedback which helped to significantly improve this manuscript.

## Appendix A

### A.1. Lateral inhibition

The processes involved in lateral inhibition are particularly relevant to the model presented in this paper. Lateral inhibition allows for competition between neurons involved in the encoding of stimuli. Along with the mechanisms of synaptic learning, this competition separates the neurons that associate the stimulus with responding, or acquisition neurons, from those which associate the stimulus with non-responding, called extinction neurons. The class of inhibitory functions that Leabra adopts are known as *k*-winners-take-all (*kWTA*) functions. A *kWTA* function ensures that no more than $k$ units out of a total of $n$ in a layer are active at any given point in time. This is attractive from a biological perspective because it captures the *set point* property of inhibitory interneurons, where the activity level is maintained through negative feedback at a roughly constant level (i.e. $k$).

#### A.1.1. kWTA function implementation

The $k$ active units in a *kWTA* function are the ones receiving the most excitatory input ($g_e$). Each unit in the layer computes a layer-wide level of inhibitory conductance ($g_i$) while updating its membrane potential such that the top $k$ units will have above threshold equilibrium membrane potentials with that value of $g_i$, while the rest will remain below firing threshold. The function computes the amount of inhibitory current $g_i^\theta$ that would put a unit just at threshold given its present level of excitatory input, where $\theta$ is the threshold membrane potential value. Computing inhibitory conductance at the threshold ($g_i^\theta$), yields

$$g_i^\theta = \frac{g_e^* g_e^- (E_e-\theta)+g_l g_l^- (E_l-\theta)}{\theta-E_i} \quad (1)$$

where $g_e^*$ represents the excitatory input minus the contribution from the bias weight and $g_l g_l^-$,$g_e g_e^-$ are the total conductances from the Potassium and Sodium channels respectively. $E_l$ and $E_e$ are the equilibrium potentials for the Potassium and Sodium channels respectively [20]. $g_i$ is computed as an intermediate value between the $g_i^\theta$ values for the $k$th and $k+1$th units as sorted by level of excitatory conductance ($g_e$). This ensures that the $k+1$th unit remains below threshold, while the $k$th unit is above it. Expressed as a formula this is given by

$$g_i = g_{k+1}^\theta + q(g_i^\theta(k)-g_i^\theta(k+1)) \quad (2)$$

where $0 < q < 1$ determines where the inhibition lies between the $k$ and $k+1$th units.

### A.2. Leabra learning algorithms

Leabra provides for a balance between Hebbian and error-driven learning. Hebbian learning is performed using a conditional principal components analysis (CPCA) algorithm. Error-driven learning is performed using GeneRec, which is a generalization of the Recirculation algorithm, and approximates Almeida–Pineda recurrent backpropagation.

#### A.2.1. Hebbian learning

The objective of the CPCA learning rule is to modify the weights for a given input unit ($x_i$) to represent the conditional probability that the input unit ($x_i$) is active when the corresponding receiving unit ($y_j$) is also active. This is expressed as

$$w_{ij} = P(x_i = 1 \,|\, y_i = 1) = P(x_i | y_j) \tag{3}$$

In Eq. (3) the weights reflect the frequency with which a given input is active across the subset of input patterns represented by the receiving unit. If an input pattern occurs frequently with such inputs, then the resulting weights from it will be relatively large. On the other hand if the input pattern occurs rarely across such input patterns then the resulting weights will be small. The following weight update rule achieves the CPCA conditional probability objective represented by Eq. (3).

$$\Delta w_{ij} = \varepsilon[y_j x_i - y_j w_{ij}] = \varepsilon y_j (x_i - w_{ij}) \tag{4}$$

where $\varepsilon$ is the learning rate parameter. The weights are adjusted to match the value of the sending unit activation $x_i$, weighted in proportion to the activation of the receiving unit ($y_j$). Thus inactivity of the receiving unit implies that no weight modification will occur. Conversely, if the receiving unit is very active (near 1), the update rule modifies the weight to match the input unit's activation. The weight will eventually come to approximate the expected value of the sending unit when the receiver is active (consistent with Eq. (3)).

#### A.2.2. Error-driven learning

GeneRec implements error backpropagation using locally available activation variables thereby making such a learning rule biologically plausible. The algorithm incorporates the notion of plus and minus activation phases. In the *minus phase*, the outputs of the network represent the expectation or response of the network, as a function of the standard activation settling process in response to a given input pattern. Then, in the *plus phase*, the environment is responsible for providing the outcome or target output activations.

The learning rule for all units in the network is given by

$$\Delta w_{ij} = \varepsilon(y_j^+ - y_j^-) x_i^- \tag{5}$$

for a receiving unit with activation $y_i$ and sending unit with activation $x_i$. The rule for adjusting the bias weights is just the same as for the regular weights, but with the sending unit activation set to 1:

$$\Delta \beta_{ij} = \varepsilon(y_j^+ - y_j^-) \tag{6}$$

The difference between the two phases of activation is an indication of the units' contribution to the overall error signal. Bidirectional connectivity allows output error to be communicated to a hidden unit in terms of the difference in its activation states during the plus and minus states ($y_j^+ - y_j^-$).

## References


[1] E.J. Kehoe, A layered network model of associative learning: learning to learn and configuration, Psychological Review 95 (4) (1988) 411–433.
[2] A.H. Klopf, A neuronal model of classical conditioning, Psychobiology 16 (2) (1988) 85–125.
[3] J.F. Medina, K.S. Garcia, M.D. Mauk, A mechanism for savings in the cerebellum, Journal of Neuroscience 21 (11) (2001) 4081–4089.
[4] M.E. Bouton, D. Swartzentruber, Analysis of the associative and occasion-setting properties of contexts participating in pavlovian discrimination, Journal of Experimental Psychology: Animal Behavior Processes 12 (4) (1986) 333–350.
[5] M.E. Bouton, D.C. Brooks, Time and context effects on performance in a pavlovian discrimination reversal, Journal of Experimental Psychology: Animal Behavior Processes 19 (2) (1993) 165–179.
[6] M.E. Bouton, Context and behavioral processes in extinction, Learning and Memory 11 (2004) 485–494.
[7] L.D. Devenport, Spontaneous recovery without interference: why remembering is adaptive, Animal Learning and Behavior 26 (2) (1998) 172–181.
[8] A. Gupta, D.C. Noelle, Lateral inhibition explains savings in conditioning and extinction, in: Proceedings of the 28th Annual Conference of the Cognitive Science Society, 2006.
[9] R.A. Rescorla, Experimental extinction, in: R.R. Mowrer, S.B. Klein (Eds.), Handbook of Contemporary Learning Theories, Erlbaum, Hillsdale, NJ, 2001, pp. 119–154.
[10] R.A. Rescorla, Retraining of extinguished pavlovian stimuli, Journal of Experimental Psychology: Animal Behavior Processes 27 (2) (2001) 115–124.
[11] R.A. Rescorla, Savings tests: separating differences in rate of learning from differences in initial levels, Journal of Experimental Psychology: Animal Behavior Processes 28 (4) (2002) 369–377.
[12] R.A. Rescorla, Comparison of the rates of associative change during acquisition and extinction, Journal of Experimental Psychology: Animal Behavior Processes 28 (4) (2002) 406–415.
[13] R.A. Rescorla, More rapid associative change with retraining than with initial training, Journal of Experimental Psychology: Animal Behavior Processes 29 (4) (2003) 251–260.
[14] R.A. Rescorla, Spontaneous recovery of excitation but not inhibition, Journal of Experimental Psychology: Animal Behavior Processes 31 (2) (2005) 277–288.
[15] A.R. Wagner, S.E. Brandon, A componential theory of pavlovian conditioning, in: R.R. Mowrer, S.B. Klein (Eds.), Handbook of Contemporary Learning Theories, Erlbaum, Hillsdale, NJ, 2001, pp. 23–64.
[16] C. Balkenius, J. Moren, Computational Models of Classical Conditioning: A Comparitive Study, Technical Report, LUCS 62. ISSN 1101-8453, 1998.
[17] G.S. Reynolds, A Primer of Operant Conditioning, Foresman and Company, Scott, 1975.
[18] R.A. Rescorla, C.D. Heth, Reinstatement of fear to an extinguished conditioned stimulus, Journal of Experimental Psychology: Animal Behavior Processes 104 (1) (1975) 88–96.
[19] J.E. Mazur, Past experience, recency, and spontaneous recovery in choice behavior, Animal Learning and Behavior 24 (1) (1996) 1–10.
[20] R.C. O'Reilly, Y. Munakata, Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain, MIT Press, 2000.
[21] A.D. Redish, S. Jensen, A. Johnson, Z. Kurth-Nelson, Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling, Psychological Review 114 (3) (2007) 784–805.
[22] S.J. Gershman, D.M. Blei, Y. Niv, Context, learning and extinction, Psychological Review 117 (1) (2010) 197–209.
[23] A.C. Courville, N.D. Daw, D.S. Touretzky, Bayesian theories of conditioning in a changing world, TRENDS in Cognitive Sciences 10 (7) (2006) 294–300.
[24] J. Pearce, G. Hall, A model for pavlovian learning: variation in the effectiveness of conditioned but not unconditioned stimuli, Psychological Review 87 (1980) 403–414.
[25] S. Grossberg, N.A. Schmajuk, Neural dynamics of attentionally modulated pavlonian conditioning: conditioned reinforcement, inhibition, and opponent processing, Psychology 15 (1987) 195–240.
[26] S. Grossberg, D. Bullock, M. Dranias, Neural dynamics of underlying impaired autonomic and conditioned responses following anygdala and orbitofrontal lesions, Behavioral Neuroscience 122 (2008) 1100–1125.
[27] S.M. Evans, Non-associative behavioral modifications in the nereis diversicolor polychate, Animal Behavior 14 (1966) 107–119.
[28] G. Morrison, J. Wen, Olfactory associative learning in caenorhabditis elegans is impaired in lrn-1 and lrn-2 mutants, Behavioral Neuroscience 113 (1999) 358–367.
[29] J.Y.M. Wen, N.N. Kumar, G. Morrison, G. Rambaldini, S. Runciman, J. Rousseau, D. van der Kooy, Mutations that prevent associative learning in c. elegans, Behavioral Neuroscience 111 (1997) 354–368.
[30] J.K. Rose, C.H. Rankin, Analysis of habituation in c. elegans, Learning Memory 8 (2001) 63–69.
[31] D.J. Felleman, D. van Essen, Distributed hierarchichal processing in the primate cerebral cortex, Cerebral Cortex 1 (1991) 1–47.
[32] E.L. White, Cortical Circuits: Synaptic Organization of the Cerebral Cortex, Stucture, Function and Theory, Birkhauser Verlag AG, 1989.

[33] R.A. Rescorla, A.R. Wagner, A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement, in: Classical Conditioning II: Current Research and Theory, Appleton-Century-Crofts, New York, 1972, pp. 64–99.
[34] S.J. Shabel, P.H. Janak, Substantial similarity in amygdala neuronal activity during conditioned appetitive and aversive emotional arousal, Proceedings of the National Academy of Sciences 106 (2009) 15031–15036.
[35] S. Maren, G.J. Quirk, Neuronal signalling of fear memory, Nature Reviews Neuroscience 5 (2004) 844–852.
[36] M.A. Morgan, J.E. LeDoux, Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats, Behavioral Neuroscience 109 (1995) 681–688.
[37] G.J. Quirk, G.K. Russo, J.L. Barron, K. Lebron, The role of ventral medial prefrontal cortex in the methods recovery of extinguished fear, Journal of Neuroscience 20 (2000) 6225–6231.
[38] M.R. Milad, G.J. Quirk, Neurons in medial prefrontal cortex signal memory for fear extinction, Nature 420 (2002) 70–74.
[39] R.F.M. Westbrook, G.M. Iordanova, R. Richardson, J.A. Harris, Reinstatement of fear to an extinguished conditioned stimulus: two roles for context, Journal of Experimental Psychology 28 (1) (2002) 97–110.
[40] M.R.M.G.J. Quirk, Neurons in medial prefrontal cortex signal memory for fear extinction, Nature 420 (2002) 70–74.
[41] G.J. Quirk, R. Garcia, F. Gonzalez-Lima, Prefrontal mechanisms in extinction of conditioned fear, Biological Psychiatry 60 (2006) 337–343.
[42] S. Maren, Long-term potentiation in the amygdala: a mechanism for emotional learning and memory, Trends in Neuroscience 23 (1999) 345–346.
[43] P.S. Bellgowan, F.J. Helmstetter, Neural systems for the expression of hypoalgesia during nonassociative fear, Behavioral Neuroscience 110 (1996) 727–736.
[44] T. Amano, C.T. Unal, D. Pare, Synaptic correlates of fear extinction in the amygdala, Nature Neuroscience 13 (2010) 1097–6256.
[45] R.S. Sutton, Learning to predict by the method of temporal differences, Machine Learning 3 (1988) 9–44.

**Ashish Gupta** has a masters degree in computer science from Vanderbilt University. He is currently working at Google as a software engineer. He enjoys applying machine learning techniques to solve everyday problems. He also enjoys working on neurocomputational models to further our understanding of how the brain works.

**Lovekesh Vig** obtained his PhD in Computer Science from Vanderbilt University, Nashville, TN. His dissertation was primarily concerned with multiple robot team or coalition formation. He is currently serving as an Assistant Professor at Jawaharlal Nehru University, New Delhi, India where he works on areas relating to Distributed Robotics, Cognitive Modeling and Machine Learning. He also holds a Masters in Computer Applications from Delhi University and an undergraduate Degree in Physics from St. Stephens College, Delhi.

**David C. Noelle** is Assistant Professor of Cognitive Science and Computer Science at the University of California, Merced. He received his Ph.D. in Cognitive Science and Computer Science from the University of California, San Diego. He completed postdoctoral training at the Center for the Neural Basis of Cognition, a joint project of Carnegie Mellon University and the University of Pittsburgh. Prior to moving to Merced, Dr. Noelle was on the faculty at Vanderbilt University. Dr. Noelle's research focuses on computational cognitive neuroscience models of the role of prefrontal cortex in learning, memory, and cognitive control.