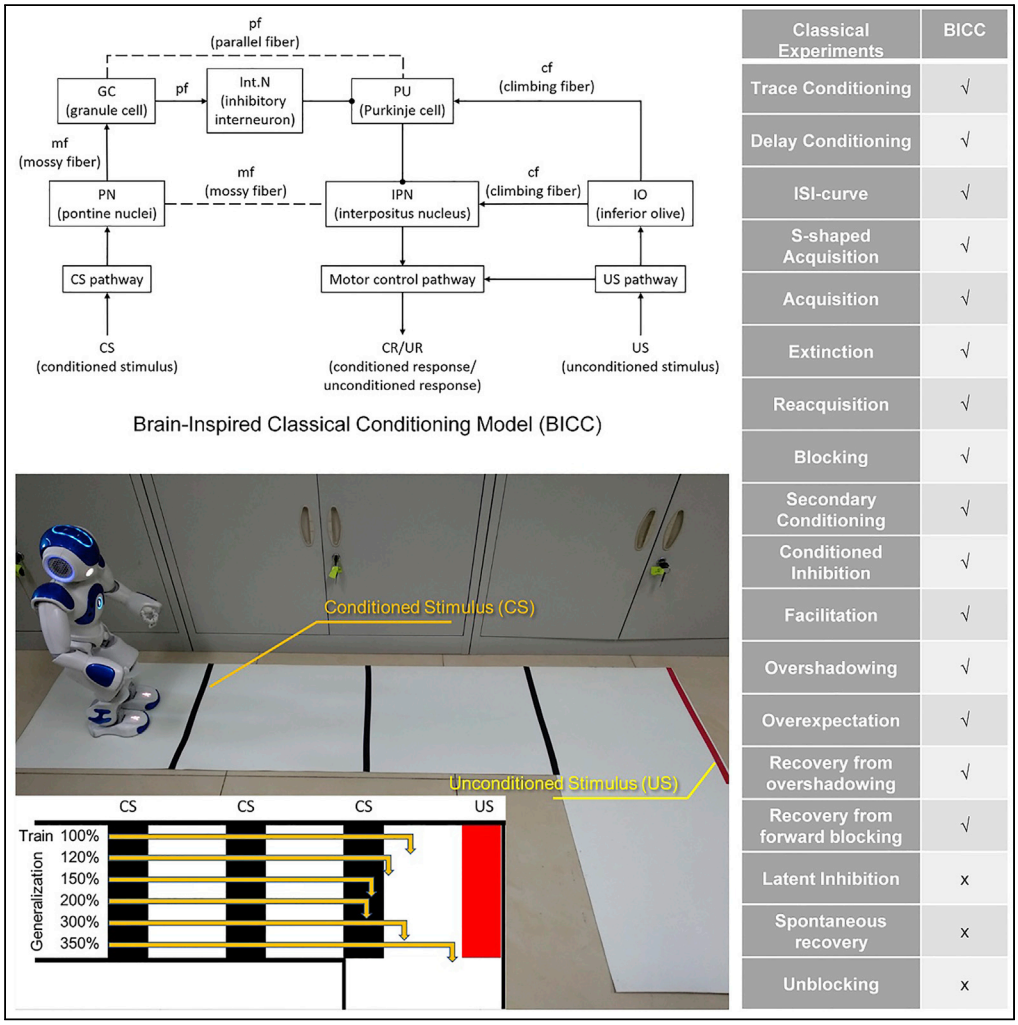
iScience

CellPress
OPEN ACCESS

Article

Brain-inspired classical conditioning model



Classical Experiments	BICC
Trace Conditioning	√
Delay Conditioning	√
ISI-curve	√
S-shaped Acquisition	√
Acquisition	√
Extinction	√
Reacquisition	√
Blocking	√
Secondary Conditioning	√
Conditioned Inhibition	√
Facilitation	√
Overshadowing	√
Overexpectation	√
Recovery from overshadowing	√
Recovery from forward blocking	√
Latent Inhibition	x
Spontaneous recovery	x
Unblocking	x

Yuxuan Zhao, Yi Zeng, Guang Qiao

yi.zeng@ia.ac.cn

HIGHLIGHTS

Classical conditioning (CC) is crucial in biological and embodied robot learning

A spiking neural network incorporates existing biological findings of CC in one model

BICC can explain a broader set of findings than other existing computational models

BICC ensures a robot gets similar biological behavior and speed generalization capability

Zhao et al., iScience 24, 101980
January 22, 2021
https://doi.org/10.1016/j.isci.2020.101980

iScience

CellPress
OPEN ACCESS

Article

Brain-inspired classical conditioning model

Yuxuan Zhao,[1,5] Yi Zeng,[1,2,3,4,5,6,*] and Guang Qiao[1]

SUMMARY

Classical conditioning plays a critical role in the learning process of biological brains, and many computational models have been built to reproduce the related classical experiments. However, these models can reproduce and explain only a limited range of typical phenomena in classical conditioning. Based on existing biological findings concerning classical conditioning, we build a brain-inspired classical conditioning (BICC) model. Compared with other computational models, our BICC model can reproduce as many as 15 classical experiments, explaining a broader set of findings than other models have, and offers better computational explainability for both the experimental phenomena and the biological mechanisms of classical conditioning. Finally, we validate our theoretical model on a humanoid robot in three classical conditioning experiments (acquisition, extinction, and reacquisition) and a speed generalization experiment, and the results show that our model is computationally feasible as a foundation for brain-inspired robot classical conditioning.

INTRODUCTION

Classical conditioning is regarded as a basic learning method for animals in which an association is built between a conditioned stimulus (*CS*) and a conditioned response (*CR*). The best-known experiment of classical conditioning was performed by Pavlov (1927). When a dog is presented with food (unconditioned stimulus, *US*), it will start to salivate (unconditioned response, *UR*). In Pavlov's research, a dog would hear a tone (*CS*) before being presented with food every time. After a number of trials, the dog started to salivate (*CR*) upon hearing a tone.

Computational model

Classical conditioning has attracted the interest of many researchers, and attempts have been made to build a computational model to reveal its mechanism. Rescorla and Wagner (1972) presented the first computational model of classical conditioning, which is named the Rescorla-Wanger model. This model can predict some important classical phenomena of classical conditioning, and it has led to a great deal of research, including modifications and alternative models. The Sutton-Barto (SB) model (Sutton and Barto, 1981) is a temporally refined extension of the Rescorla-Wanger model. This model learns to increase its response rate in anticipation of increased stimulation, and it can account for more phenomena observed in classical conditioning than the Rescorla-Wanger model can; furthermore, it has served as the precursor of many later models. The temporal difference (TD) model (Sutton and Barto, 1987) is an extension of the SB model. This model takes the form of a temporal difference prediction method and can successfully model the inter-stimulus interval (ISI) effect, which is regarded as a primary real-time effect of classical conditioning. Harry Klopf (1988) proposed a neuronal model that is modified from the Hebbian model to be more consistent with animal learning phenomena; this model can predict a wide range of classical conditioning phenomena. Schmajuk and DiCarlo (1992) presented a multilayer neural network called the S-D model. They mapped the nodes and connections onto regional cerebellar, cortical, and hippocampal circuits to obtain a model that can correctly describe the effects of hippocampal and cortical lesions on conditioning. Balkenius and Moren (1999) described a computational model of classical conditioning that is built on the assumption that the goal of learning is the prediction of a temporally discounted reward or punishment based on the current stimulus situation; notably, this model is well suited for robotic implementation. Johansson and Lansner (2002) presented an associative model of classical conditioning that is composed of a number of interconnected Bayesian confidence propagation neural networks (BCPNNs) implemented on the basis of Hebbian learning, and the output of this model fits the results of classical conditioning experiments. Zuo et al. (2005) introduced a spiking-neuron-based cognitive model. This model can simulate the

[1]Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

[2]Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

[3]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

[4]School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100190, China

[5]These authors contributed equally

[6]Lead contact

*Correspondence: yi.zeng@ia.ac.cn

https://doi.org/10.1016/j.isci.2020.101980

learning process of classical conditioning with a reflex arc structure and a reinforcement learning method based on the Hebb rule, and an inverted pendulum experiment validated the effectiveness of this model. Liu et al. (2008) built a model with classical conditioning behaviors based on a Bayesian network (CRMBBN). This model constructs cause-effect relationships between classical sensing and nonclassical conditional sensing by means of a Bayesian network, and it can successfully simulate many phenomena, such as acquisition, inter-stimulus effects, and extinction. Liu and Ding (2008) presented a dynamic policy adaptation framework inspired by classical conditioning. This model can successfully realize the self-learning process of classical conditioning and achieve adaptive network policy management. Antonietti et al. (2017) developed a detailed spiking cerebellar microcircuit model that can reproduce eyeblink classical conditioning and successfully fits real experimental datasets from humans.

Here, we build a brain-inspired classical conditioning (BICC) model that integrates and adopts existing biological findings of classical conditioning. With a quaternionic-rate-based synaptic learning rule, which is equivalent to spike-timing-dependent plasticity (STDP) (Bi and Poo, 2001) on a timescale of seconds, the BICC model could predict the majority of classical conditioning phenomena. The computational model of biological classical conditioning enables a robot gets similar learning behavior and the capability of speed generalization. Furthermore, the changes in synaptic weights in this model may hint at the biological mechanism of classical conditioning.

Classical experiments

There are several classical conditioning experiments that can be used to verify the effectiveness of computational models. To enable reasonable comparisons with other well-known computational models, we follow the classical conditioning experiments outlined by Balkenius and Moren (1998). The stimulus before the + presents first, and the stimulus after the + presents later. The parenthesis indicates that the stimuli are presented and end simultaneously. The ⇒ indicates the result of training. The → indicates the result of the stimulus.

Acquisition

Acquisition is the ability to establish an association between a stimulus and a response, and it is the most fundamental process in classical conditioning. In an acquisition experiment, a *CS* is presented first and a *US* is presented subsequently after a small time interval for several trials; then, the response will be induced when the *CS* is presented on its own. This acquisition progress can be described as follows. For an acquisition experiment involving an eyelid response in the albino rabbit, the response level forms an S-shaped curve similar to a sigmoid function (Balkenius and Moren, 1998; Schneiderman et al., 1962).

$$CS + US \Rightarrow CS \rightarrow R$$

Inter-stimulus interval effect

The ISI effect is a primary real-time effect of classical conditioning (Sutton, 1990). The ISI represents the time interval between the presentation of the *CS* and *US*, and it can be divided into three types (Balkenius and Moren 1999): delay conditioning A, delay conditioning B, and trace conditioning (Figure S1). In delay conditioning A, the *US* is presented immediately when the *CS* terminates. In delay conditioning B, the *CS* is still present when the *US* is presented, and the CS and US terminate simultaneously. In trace conditioning, the CS and US have fixed lengths, and the CS terminates before the onset of the US. In empirical studies conducted by Schneiderman and Gormezano (1964) and Smith et al. (1969), the ISI-CR frequency function revealed a concave-down shape during the acquisition and extinction process.

Extinction

In an extinction experiment, the acquired response will disappear gradually if only a *CS* is presented without the subsequent *US*. The extinction process can be described as follows:

$$CR \Rightarrow CS \rightarrow no\ R$$

Reacquisition effect

The reacquisition effect is demonstrated when an animal relearns a previously extinguished association, and the relearning phase is faster than the initial learning phase.

Blocking

Blocking refers to the following phenomenon: when a first stimulus CS_1 has been associated with a response and a second stimulus CS_2 then is presented and ends simultaneously with CS_1, the attempt to associate the second stimulus CS_2 with the response will be blocked. Blocking experiments show that the association of a stimulus with a response is not independent of earlier learning. The blocking process can be described as follows, where the parentheses are used to indicate that CS_1 and CS_2 are presented and end simultaneously.

$$CS_1 + US \Rightarrow CS_1 \rightarrow R$$
$$(CS_1 + CS_2) + US \Rightarrow CS_2 \rightarrow no\ R$$

Secondary conditioning

In a secondary conditioning experiment, CS_1 has been associated with the response induced by the *US*, and CS_1 is then used as the *US* for CS_2 to build an association to the response. The effect of such secondary conditioning is typically weak, and CS_1 will undergo extinction, whereas CS_2 is reinforced. The secondary conditioning progress can be described as follows:

$$CS_1 + US \Rightarrow CS_1 \rightarrow R$$
$$CS_2 + CS_1 \Rightarrow CS_2 \rightarrow R$$

Conditioned inhibition

In a conditioned inhibition experiment, CS_2 and CS_0 have been associated with the response, and then a third stimulus CS_1 is presented and ends simultaneously with one of the previously conditioned stimuli CS_0 without the *US*. In the test phase, CS_1 will inhibit the ability of CS_2 to induce the response. The conditioned inhibition process can be described as follows, where the parentheses are used to indicate that the stimuli are presented and end simultaneously.

PhaseI	$CS_2 + US$
PhaseII	$CS_0 + US$
	$(CS_0 + CS_1) + no\ US$
Test	$(CS_1 + CS_2) \rightarrow no\ R$

Facilitation by an intermittent stimulus

Under normal acquisition conditions, the conditioning to CS_1 is weak in the case of trace conditioning due to the long ISI. Under conditions of facilitation, the conditioning to CS_1 is facilitated by an additional CS_2. The facilitation process can be described as follows:

Normal	$CS_1 + US \Rightarrow CS_1 \rightarrow weak\ R$
Facilitated	$CS_1 + CS_2 + US \Rightarrow CS_1 \rightarrow strong\ R$

Overshadowing

The CS_1 and CS_2 are presented and ended simultaneously under the conditions of overshadowing; the associative strength acquired by CS_1 and CS_2 are weaker than the CS_1 or CS_2 conditioned alone in the normal acquisition condition (Angulo et al., 2020). The overshadowing process can be described as follows:

Normal	$CS_1 + US \Rightarrow CS_1 \rightarrow strong\ R$
	$CS_2 + US \Rightarrow CS_1 \rightarrow strong\ R$
Overshadowing	$(CS_1 + CS_2) + US \Rightarrow CS_1 \rightarrow weak\ R$
	$(CS_1 + CS_2) + US \Rightarrow CS_2 \rightarrow weak\ R$

Overexpectation

The CS_1 and CS_2 have been associated with the response, respectively, then the following CS_1-$CS2$ presentations result in a weight decrement (Rescorla and Wagner, 1972). The overexpectation process can be described as follows:

PhaseI	$CS_1 + US \Rightarrow CS_1 \rightarrow R$
PhaseII	$CS_2 + US \Rightarrow CS_2 \rightarrow R$
Test	$(CS_1 + CS_2) + US \Rightarrow weight\ decrement$

Recovery from overshadowing

In the overshadowing experiment, the extinction of the CS_1 will lead to an increased responding to the CS_2 (Matzel et al., 1985). The recovery from overshadowing process can be described as follows:

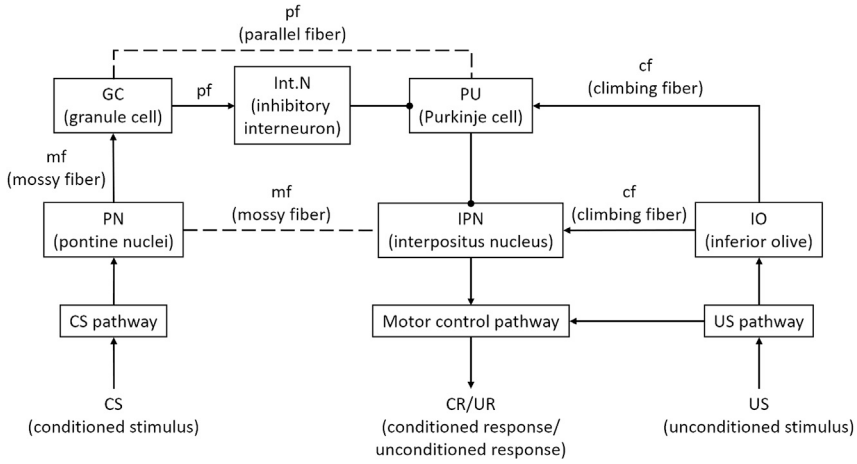


Figure 1. The architecture of the BICC model

The arrows and dots represent excitatory and inhibitory synapses, respectively, and the dotted lines represent excitatory or inhibitory synapses depending on the results of synaptic plasticity. The CS pathway and the US pathway are used for recognizing the *CS* and *US* through traditional pattern recognition methods and for transferring the information to the *PN* and *IO*, respectively. The *PN* projects the information on the *CS* to the *IPN* and *GC* via an *mf*. The *IO* projects the information on the *US* to the *IPN* and *PU* via a *cf*. The *GC* transfers the stimulus from the *PN* to the *PU* and *Int.N* via a *pf*. The *PU* receives inhibitory stimulation from the *Int.N*, excitatory stimulation from the *IO* via a *cf*, and stimulation from the *GC* via a *pf*. The *IPN* receives inhibitory stimulation from the *PU* and excitatory stimulation from the *IO* via a *cf* and from the *PN* via an *mf*. The motor control pathway receives excitatory stimulation from the *IPN* if the *IPN* is activated and then induces the *CR*, or it receives excitatory stimulation from the US pathway and then induces the *UR*.

$$(CS_1 + CS_2) + US \Rightarrow (CS_1 + CS_2) \rightarrow weak\ R$$
$$CS_2 + US \Rightarrow CS_2 \rightarrow strong\ R$$

Recovery from forward blocking

In the forward blocking experiment, the extinction of the blocker CS_1 will lead to an increased response to the blocked CS_2 (PinenO et al., 2005). The recovery from forward blocking process can be described as follows:

$$CS_1 + US \Rightarrow CS_1 \rightarrow R$$
$$(CS_1 + CS_2) + US \Rightarrow CS_2 \rightarrow no\ R$$
$$CS_2 + US \Rightarrow CS_2 \rightarrow R$$

RESULTS

The neural circuit underlying delay eyeblink conditioning has been well described in (Hansel et al., 2001; Ten Brinke et al., 2019; Wang et al., 2018; Hogri et al., 2015; Takehara-Nishiuchi, 2018), and we propose our BICC model based on these findings. The architecture of the BICC model is shown in Figure 1.

Model evaluation

In this section, we use the changes in synaptic weight between the PN (pontine nuclei) and the IPN (interpositus nucleus) to evaluate this model; NP_{CS_i} represents the neuron population of the corresponding *CS*in the PN, and NP_R represents the neuron population that controls the response in the IPN. The PN deliver the information from the *CS*, and the IPN generates the *CR* via the motor control pathway, as introduced in the Methods section.

Inter-stimulus interval effects

We use both delay and trace conditioning experiments to test our model, and the variation in the synaptic weight between NP_{CS} and NP_R is shown in Figure 2A. The curves initially show a marked increase and then exhibit a concave-down shape, which is consistent with the results of the rabbit experiment presented in Schneiderman and Gormezano (1964) and Smith et al. (1969). There is an optimal interval time for learning under every condition in this model, being 1.3 s for trace conditioning and 3.1 s for both types of delay conditioning.

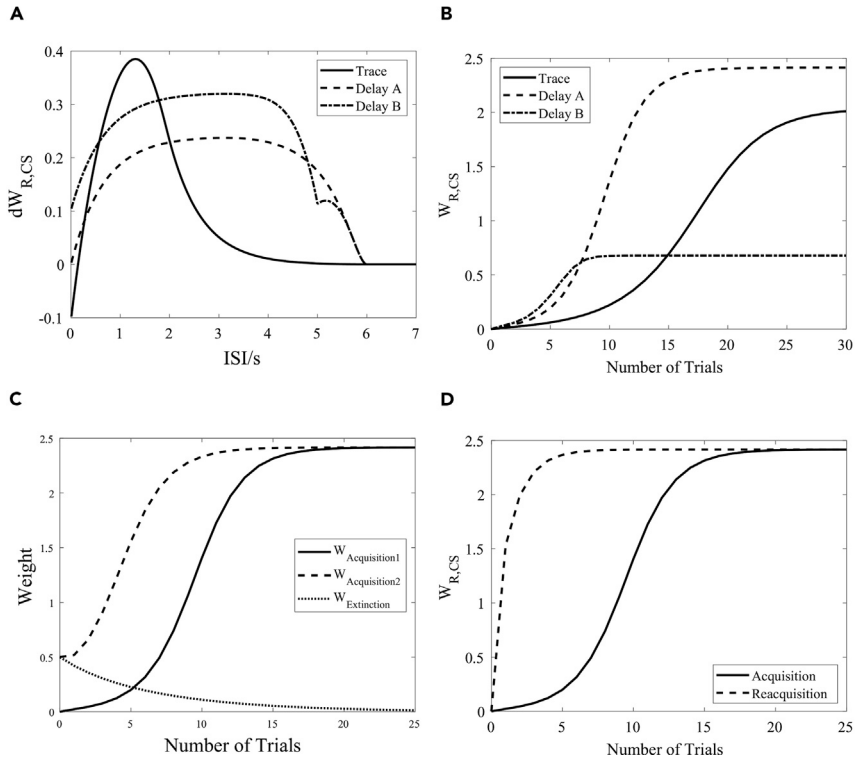


Figure 2. The results of inter-stimulus interval effects, learning curves, acquisition and extinction, and reacquisition experiments

(A) The inter-stimulus interval effects in delay and trace conditioning experiments. (1) In the delay conditioning A, the duration of the *CS* varies from 0 to 6 s, and then the *US* is presented immediately after the *CS* and continues for 2 s. The ISI is equal to the length of the *CS*. (2) In the delay conditioning B, the duration of the CS varies from 1 to 7 s, whereas the *US* continues for 1 s, and the *CS* and *US* end simultaneously. The ISI is equal to the difference between the length of the *CS* and the length of the *US*. (3) In trace conditioning, the *CS* and *US* each continue for 2 s. The ISI is equal to the difference between the start time of the *CS* and the start time of the *US*, and it varies from 0 to 6 s.

(B) The learning curves in the model. (1) In delay conditioning A, the *CS* and *US* each continue for 2 s, so the interval time is 2 s. (2) In delay conditioning B, the *CS* continues for 3 s and the *US* continues for 1 s, so the interval time is 2 s. (3) In trace conditioning, the *CS* is presented at 0 s and continues for 2 s, and the *US* is presented at 2.5 s and continues for 2 s, so the interval time is 2.5 s.

(C) The variations in synaptic weight between NP_{CS} and NP_R. $W_{Acquisition1}$ and $W_{Acquisition2}$ represent the weight variations in the acquisition experiment when the initial weight is 0 and 0.5, respectively. $W_{Extinction}$ represents the weight variations in the extinction experiment, where there is only a *CS*.

(D) Reacquisition experiment. It is easy to see that fewer trials are needed in the reacquisition condition (fewer than 10 trials) than in the acquisition condition (approximately 25 trials) to achieve the same weight.

Learning curves

Figure 2B shows the results for the learning curves in the model. Under each condition, the curve is an S-shaped acquisition curve. With fewer than eight trials, the synaptic weight between NP_{CS} and NP_R is greater in the case of delay conditioning B than in the case of delay conditioning A, and the smallest weight is observed in the trace conditioning case. As the number of trials increases, the synaptic weight increases to a stable constant; ultimately, the synaptic weight under delay conditioning A is greater than that under trace conditioning, and the smallest final weight is observed in the case of delay conditioning B. We therefore select delay conditioning A for the remainder of the experiments unless otherwise stated, with both the *CS* and the *US* continuing for 2 *s*.

Acquisition, extinction, and reacquisition experiments

Figure 2C shows the variations in synaptic weight between NP_{CS} and NP_R in the acquisition and extinction experiments. In the acquisition experiment, the *CS* is presented at 0 s and ends at 2 s and the *US* is presented at 2 s and ends at 4 s. In the extinction experiment, only a *CS* is presented at 0 s and ends at 2 s, without a *US*. The results of the reacquisition experiment are shown in Figure 2D. It is easy to see that in the reacquisition experiment, fewer

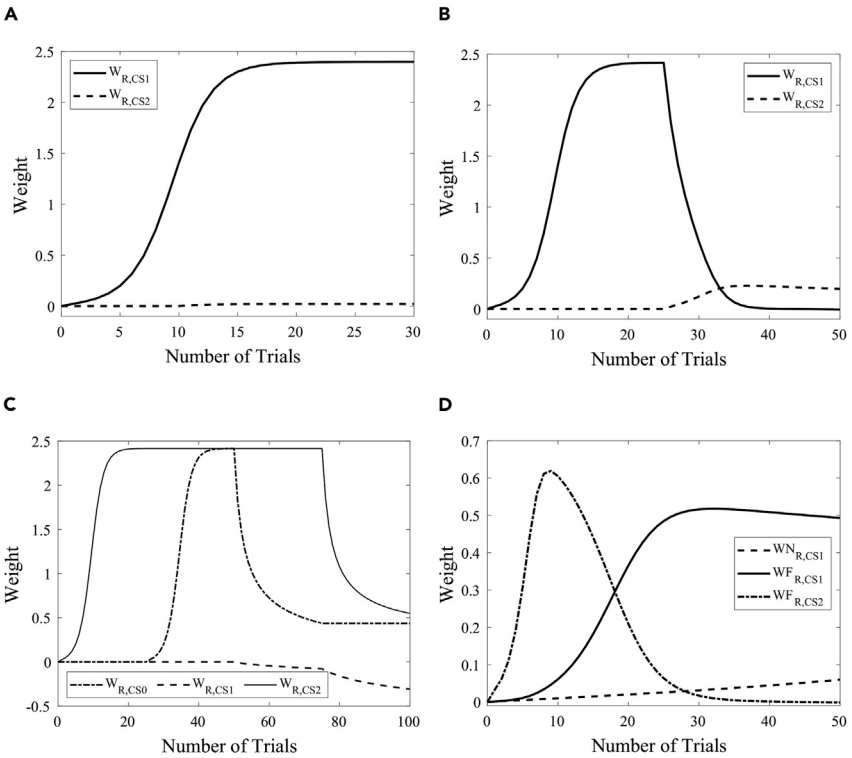


Figure 3. The results of blocking, secondary conditioning, conditioned inhibition, and facilitation experiments

(A) Blocking experiment. In the first 10 trials, only CS_1 and *US* are presented to build a conditioned response. In the remainder of the trials, CS_1 and CS_2 are presented and end simultaneously, and then the *US* is presented subsequently.

(B) Secondary conditioning experiment. In the first 25 trials, CS_1 is presented first and *US* is presented subsequently, as in the normal acquisition experiment. In the last 25 trials, CS_1 is treated as the unconditioned stimulus and is presented after CS_2.

(C) Conditioned inhibition experiment. In the first 25 and second set of 25 trials, CS_2 and CS_0, respectively, are combined with *US* for conditioning on the response. In the subsequent 25 trials, CS_0 and CS_1 are presented and end simultaneously, without *US*. In the last 25 trials, only CS_1 and CS_2 are presented.

(D) Facilitation experiment. CS_1 continues for 2 *s*, CS_2 continues for 3 *s*, and the *US* continues for 1 *s*. During normal acquisition, CS_1 is presented first, and 2 *s* later, the *US* is presented. During facilitated acquisition, CS_1 is presented first, CS_2 is presented immediately when CS_1 ends, and 2 *s* after CS_1 ends, the *US* is presented. CS_2 and the *US* end simultaneously.

trials are required to achieve a higher synaptic weight between NP_{CS} and NP_R than in the acquisition experiment. This is because during acquisition or extinction, not only the synaptic weight but also the number of synapses involved changes. In the reacquisition experiment, more synapses are involved in the synaptic weight changes, so the learning rate is faster than in the acquisition stage. We combine the acquisition, extinction, and reacquisition experiments into a single overall experiment. The results are shown in Figure S2.

Blocking experiment

Figure 3A shows the results of the blocking experiment. It is easy to see that the synaptic weight between NP_{CS_2} and NP_R is too small to induce a response. In the blocking stage, the changes in single synaptic weights between NP_{CS_1} and NP_R and between NP_{CS_2} and NP_R are identical because of the synchronization of CS_1 and CS_2, but there are many more synapses involved in NP_{CS_1} than in NP_{CS_2}, which causes the change in W_{R,CS_1} to be much greater than that in W_{R,CS_2}.

Secondary conditioning experiment

Figure 3B shows the results of the secondary conditioning experiment. Here, CS_1 is treated as the *US* to build an association between CS_2 and the response, and the corresponding synaptic weight is typically weak because the synaptic weight between NP_{CS1} and NP_R exhibits an extinction effect at the same time.

Conditioned inhibition experiment

Figure 3C shows the results of the conditioned inhibition experiment. In the first and second sets of 25 trials, excitatory synapse connections to NP_R are built for NP_{CS_2} and NP_{CS_0}, respectively. In the third set of 25 trials, NP_{CS_0} exhibits an extinction effect, whereasNP_{CS_1} builds inhibitory synapse connections to NP_R because of synchronism. In the last 25 trials, the number of synapses between NP_{CS_1} and NP_R increase because of the inhibitory connections and the negative weight changes. At the beginning of the last set of 25 trials, the inhibition effect from CS_1 is not sufficient to completely inhibit the excitatory input from CS_2, so CS_2 can still induce the response. With the enhancement of the inhibition effect from CS_1, at the end of the last set of 25 trials, CS_1 can inhibit the response induced by CS_2. In the extinction experiment, more than 50 trials are needed for CS_2 to lose the ability to induce the response, whereas in the conditioned inhibition experiment, fewer than 25 trials are needed because of the inhibition effect from CS_1.

Facilitation experiment

Figure 3D shows the results of the facilitation experiment. It is easy to see that the synaptic weight between NP_{CS_1} and NP_R is stronger under facilitated acquisition than under normal acquisition. Under normal acquisition, the synaptic weight is weak because of the long ISI. Under facilitated acquisition, CS_2 is conditioned on the response, and the response will be induced twice, by CS_2 and the *US*, thus causing CS_1 to build stronger synaptic connections.

Overshadowing experiment

Figure 4A shows the results of the overshadowing experiment. The synaptic weights between NP_{CS_1} and NP_R and NP_{CS_2} and NP_R are identical because of the synchronization of CS_1 and CS_2. It is easy to see that the synaptic weight in the overshadowing condition is weaker than that in the normal acquisition condition. In the overshadowing condition, the NP_R is stimulated by CS_1 and CS_2 simultaneously, and they contribute equally to build an association to response. So, the synaptic weight in the overshadowing condition is weaker, and it is about half of that in the normal acquisition condition.

Overexpectation experiment

Figure 4B shows the results of the overexpectation experiment. The synaptic weight between NP_{CS_1} and NP_R and NP_{CS_2} and NP_R have decreased in the last 25 trials, and it provided that the following CS_1-*CS2* presentations result in a weight decrement. In the first and subsequent 25 trials, the CS_1 and CS_2 build a *CR* with US, respectively. In the last 25 trials, the CS_1 and CS_2 stimulate the NP_R simultaneously. So, the firing rate of the response neuron increases faster and lasts longer, and it makes an extinction effect until the model is stable again.

Recovery from overshadowing experiment

Figure 4C shows the results of the recovery from overshadowing experiment. The synaptic weight between NP_{CS_2} and NP_R has increased when the CS_1 ended, and it provided that the extinction of the CS_1 will lead to an increased respondse to the CS_2. The NP_R is only stimulated by CS_2 when the CS_1 ended, and there is an acquisition effect when the *US* is presented.

Recovery from forward blocking experiment

Figure 4D shows the results of the recovery from forward blocking experiment. The synaptic weight between NP_{CS_2} and NP_R has increased when the CS_1 ended, and it provided that the extinction of the blocker CS_1 will lead to an increased response to the blocked CS_2. Similar to the recovery from overshadowing experiment, the NP_R is only stimulated by CS_2 when the CS_1 ended, and there is an acquisition effect when the *US* is presented.

The results of the presented model in the various experiments and the comparison results with existing models are summarized in Table 1.

Robotic classical conditioning experiments

We use acquisition, extinction, and reacquisition experiments and speed generalization experiment to evaluate this model on a humanoid robot.

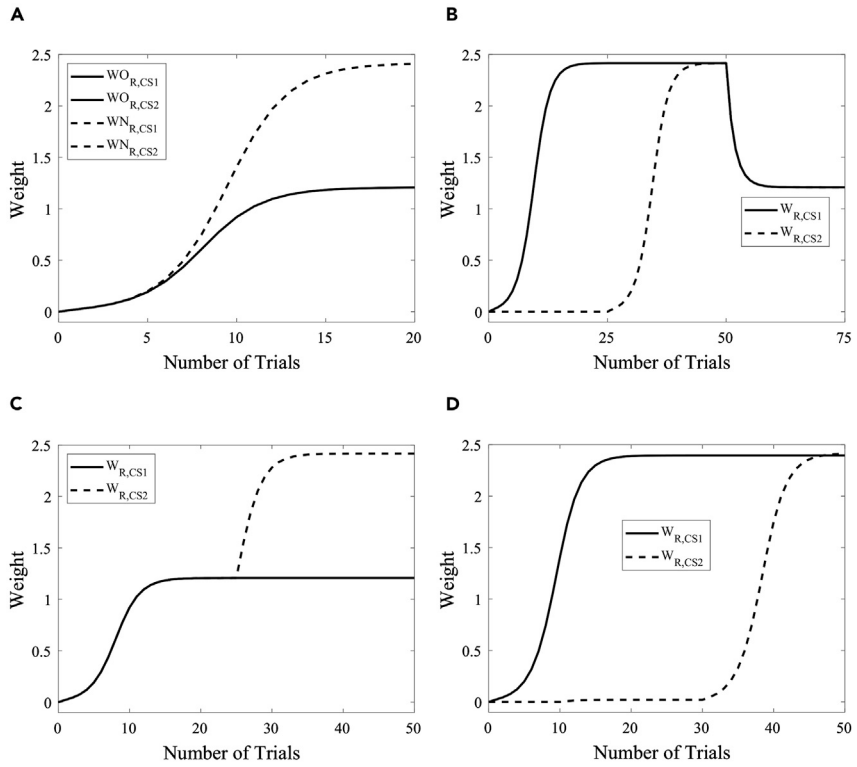


Figure 4. The results of overshadowing, overexpectation, recovery from overshadowing, and recovery from forward blocking experiments

(A) Overshadowing experiment. In the normal acquisition condition, the CS_1 and CS_2 build a conditioned response with the *US*. In the overshadowing condition, the CS_1 and CS_2 are presented and end simultaneously and then the *US* is presented subsequently. The dotted line shows the weight changing in the normal acquisition condition. The solid line shows the weight changing in the overshadowing condition.

(B) Overexpectation experiment. In the first 25 trials, the CS_1 is presented first and the *US* is presented subsequently. In the subsequent 25 trials, the CS_2 is presented first and the *US* is presented subsequently. In the last 25 trials, the CS_1 and CS_2 are presented and end simultaneously, and then the *US* is presented subsequently.

(C) Recovery from overshadowing experiment. The first 25 trials are the overshadowing process, the CS_1 and CS_2 are presented and end simultaneously, and then the *US* is presented subsequently. The last 25 trials are the recovery process, the CS_1 is ended, and the CS_2 presented first and the *US* is presented subsequently.

(D) Recovery from forward blocking experiment. The first 30 trials are the blocking process. In the first 10 trials, the CS_1 is presented first and the *US* is presented subsequently. In the subsequent 20 trials, the CS_1 and CS_2 are presented and end simultaneously, and then the *US* is presented subsequently. The last 20 trials are the recovery process, wherein the CS_1 is ended, and the CS_2 presented first and the *US* is presented subsequently.

Acquisition, extinction, and reacquisition experiments

We selected three classical conditioning experiments—acquisition, extinction, and reacquisition—to evaluate this model on a humanoid robot. A red robot was used as the participant robot, and the various stimuli used in the experiments are shown in Figure 5A. We use template matching to identify different stimuli. The sequences of stimuli in these experiments are shown in Figure 5B. In the acquisition experiment, the participant robot first was shown a red fist toy (*CS*), then was shown a blue robot (*US*), and subsequently took an avoidance action (*UR*). After learning, the participant robot would take an avoidance action (*CR*) when it saw the red fist toy (*CS*). In the extinction experiment, only the *CS*was presented; after several trials, the participant robot did not perform the *CR* upon seeing the red fist toy (*CS*). In the reacquisition experiment, the participant robot could establish the *CR* in fewer trials than in the acquisition experiment. The experimental results are shown in Figures 5C and 5D.

Speed generalization experiment

The speed generalization experiment is that the robot trained at a slow speed on a navigation task; then it could navigate the track at a higher speed without training (Herreros et al., 2013). The humanoid robot Nao is a biped

Table 1. Comparison results with existing models

Classical experiments	SB	TD	Klopf	SD	Balkenius	BCPNN	CRMMBBN	Liu	BICC
Trace conditioning	*	*	*	*	*	*	–	–	*
Delay conditioning	–	*	o	*	*	*	–	–	*
ISI curve	–	*	o	*	o	o	–	–	*
S-shaped acquisition	–	–	*	–	*	*	–	–	*
Acquisition	*	*	*	*	*	*	*	*	*
Extinction	*	*	*	*	*	*	*	*	*
Reacquisition	–	–	o	*	–	–	*	*	*
Blocking	*	*	*	*	*	*	–	*	*
Secondary conditioning	o	o	*	–	*	*	–	*	*
Conditioned inhibition	*	*	*	*	*	*	–	–	*
Facilitation	*	*	*	*	*	o	–	–	*
Overshadowing	–	–	*	*	*	–	–	–	*
Overexpectation	–	–	*	–	–	–	–	–	*
Recovery from overshadowing	–	–	–	–	–	–	–	–	*
Recovery from forward blocking	–	–	–	–	–	–	–	–	*
Latent inhibition	–	–	–	–	–	–	–	–	–
Spontaneous recovery	–	–	–	–	–	–	–	–	–
Unblocking	–	–	–	–	–	–	–	–	–

This table is adapted from Johansson and Lansner (2002) and Balkenius and Moren (1998). An * means that the model could reproduce the correlation feature, o means that the model could reproduce partially, and - means that it is not mentioned or was unable to reproduce.
SB (Sutton and Barto 1981), TD (Sutton and Barto 1987), Klopf (Harry Klopf 1988), SD (Schmajuk and DiCarlo 1992), Balkenius (Balkenius and Moren 1999), BCPNN (Johansson and Lansner 2002), CRMBBN (Liu et al., 2008), Liu (Liu and Ding 2008).

robot. According to its documentation, the speed parameter is the fraction of the maximum speed, such as 100% means the full speed. According to our verification, may be due to the aging problem, the robot cannot move accurately at a given speed, and often deviates from the direction when walking in a straight line. So, we test the speed generalization capabilities of the model in both real and simulation environments.

The real environment is shown as Figure 6A. In the real environment, the robot is trained at 50% speed and tested at 100% speed. We use simple image recognition algorithms to recognize the CS (black line, Figure 6B) and the US (red line, Figure 6C). For the black line recognition, we retain the central area of the image, then convert the image to a binary image according to the given threshold value, delete the small-area object, and finally perform horizontal line detection to complete the recognition. For the red line recognition, we extract the matching color area by setting the threshold value and then detect the horizontal line to complete the recognition. The result of the real environment is shown as Figure 7.

The simulation environment is shown as Figure 8A. In the simulation environment, we use two experiments to test the speed generalization capacity of the model. In the first experiment, we test the induced time of CR and UR at a given speed in the range of 100%–200% with 10% speed increments. The result is shown as Figure 8B. In the second experiment, we test the induced distance of CR at a given speed at 100%, 120%, 150%, 200%, 300%, 350%, and 400%. The result is shown as Figure 8C. And if the speed is greater than 350%, the robot experiment fails.

The results of real and simulation environments show that the proposed model has the capacity of speed generalization.

DISCUSSION

The BICC model exhibits long-term depression (LTD) at GC (granule cell)-PU (Purkinje cell) synapses and long-term potentiation (LTP) at PN-IPN synapses, which is consistent with the findings from electrophysiological experiments on eyeblink conditioning presented in Koekkoek et al. (2003) and Steuber et al.

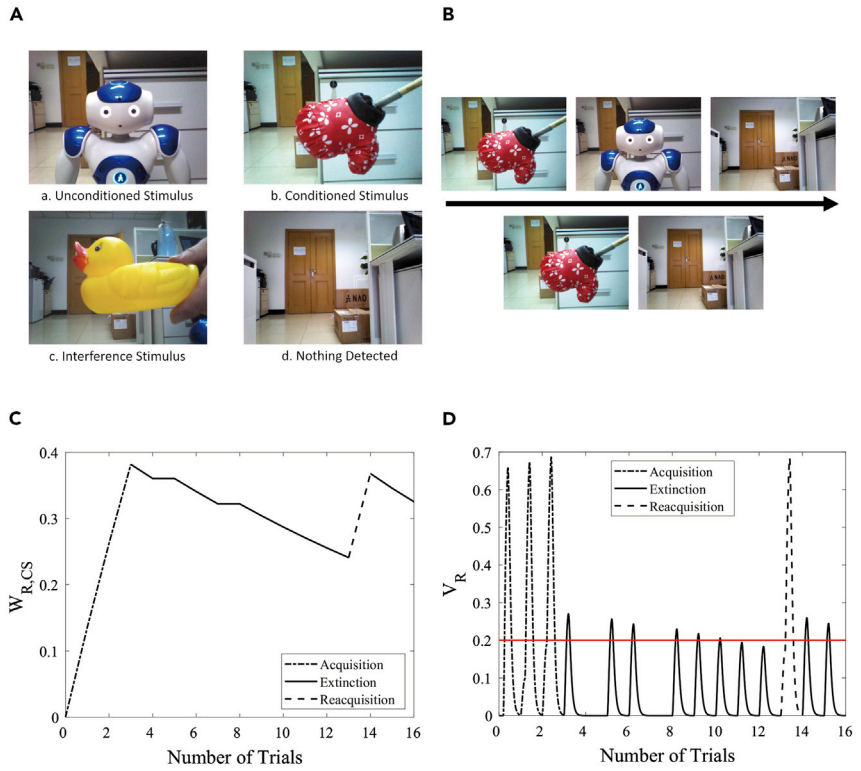


Figure 5. The acquisition, extinction, and reacquisition experiments on humanoid robot

(A) The stimuli used in the experiments. (a) The blue robot is an unconditioned stimulus that can be regarded as the participant robot's natural enemy; therefore, the participant robot will take an avoidance action upon seeing the blue robot. (b) The red fist toy is a conditioned stimulus. (c) The yellow duck is an interference stimulus used to prove that only the learned stimulus can induce the response. (d) Nothing detected means that there is no stimulus.

(B) The sequences of stimuli for acquisition (upper) and extinction (below).

(C and D) (C) Dynamic changes of weight in robotic classical conditioning experiments. (D) Dynamic changes of response in robotic classical conditioning experiments. The red line indicates the response threshold. The total number of trials is 16, consisting of 3 acquisition trials, 8 extinction trials, 1 reacquisition trial, and 2 extinction trials in addition to 2 interference stimuli (in trials 4 and 7) in the first extinction experiment.

(2007) and Pugh and Raman, (2006, 2008), respectively. LTD is also exhibited at PN-IPN synapses in this model, representing another mechanism of synaptic plasticity that is consistent with the electrophysiological experiments reported in Zhang and Linden (2006), but there have been fewer reports of this mechanism than of the former.

LTD at GC-PU synapses

If only the *CS* is presented before learning, the GC (granule cell) receives the *CS* from the PN and then projects it to an Int.N (inhibitory interneuron). The inhibitory input from the Int.N will inhibit the excitatory input from the GC and the spontaneous firing of the PU, so the firing rate of the PU will immediately drop to zero; in other words, the firing of the PU will be paused. The synaptic weight between the GC and PU will not change because the postsynaptic neurons in the PU is not fired.

If only the *US* is presented before learning, the *US* is projected to the motor control pathway and induces the *UR*. In parallel, the IO (inferior olive) receives the *US* from the US pathway and then projects it to the IPN and PU via the cf (climbing fiber). The excitatory input from the IO to the PU will enhance the inhibition from the PU to the IPN. The synaptic weight between the GC and PU remains unchanged because the presynaptic neuron in the GC is not fired.

In the acquisition experiment, the firing rate of the PU will gradually decay to zero because of the additional excitatory input from the *US*. With the increased firing rate of the presynaptic neurons in the GC and the

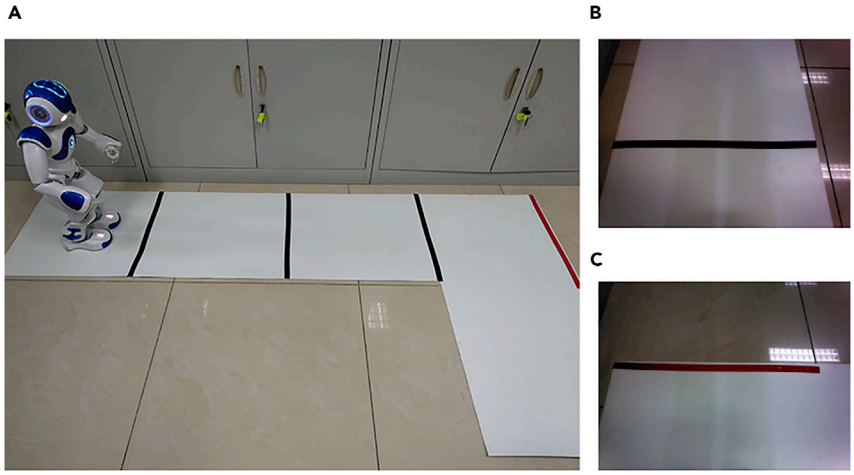


Figure 6. Speed generalization experiment in real environment

(A) The real environment. The white runway is the track for the robot to navigate. The red line is the US, which means the robot needs to turn right to avoid leaving the runway, and the black line is the CS.

(B) The CS perceived in robot vision.

(C) The US perceived in robot vision.

decreasing firing rate of the postsynaptic neurons in the PU, the synaptic weight between the GC and PU will decrease and exhibit LTD. An intuitive explanation of how spike-based STDP can influence synaptic efficiency through a rate-based mechanism can be found in Bengio et al. (2017). In short, the synaptic weight is updated in proportion to the product of the presynaptic firing rate and the temporal rate of change in activity on the postsynaptic side. The synaptic weight is updated in our model based on more factors, as well, as detailed in Qiao et al. (2017). The synaptic weight between the GC and PU will decrease with repeated pairings of the *CS* and *US*, whereas the inhibitory input from the Int.N and the excitatory input from the IO remain unchanged, so the Purkinje cell will pause spontaneous firing. This phenomenon is consistent with the electrophysiological experiments on eyeblink conditioning presented in Wetmore et al. (2014) and Hansel et al. (2001).

LTP and LTD at PN-IPN synapses

The single synaptic weight changes in a single trial in the acquisition and extinction experiments are shown in Figure 9A, and the firing rates of the neurons in the acquisition and extinction experiments are shown in Figures 9B and 9C, respectively. In the acquisition experiment, the change in the synaptic weight is negative because the temporal rate of change of the postsynaptic neurons in the IPN, which is represented as V_R in Figures 9B, is smaller than that of the presynaptic neuron in the PN, which is represented as V_{CS} in Figure 9B, from 0 to 2 s. The *CS* ends at 2 s; then the *US* is presented, continues for 2 s, and finally ends at 4 s. From 2 to 4 s, the firing rate of the presynaptic neuron decreases because the *CS* has ended, whereas the firing rate of the postsynaptic neuron increases because the *US* is being presented. The change in the synaptic weight is positive because of the decreasing firing rate of the presynaptic neuron and the increased firing rate of the postsynaptic neuron. This model exhibits the acquisition effect if the positive term is greater than the negative term and exhibits the extinction effect if the positive term is less than the negative term, and the model reaches a steady state when the positive term is equal to the negative term.

In the acquisition experiment, the number of excitatory synapses between the PN and IPN increases, which is consistent with the electrophysiological experiments on eyeblink conditioning presented in Kleim et al. (2002) and Weeks et al. (2007).

Our model suggests that the cerebellar cortex, especially the IPN, plays a critical role in classical conditioning. In our model, the LTD at GC-PU synapses leads to a reduction in the excitatory input from the GC. Although there is an excitatory input from the IO when the *US* is presented, the PU will be paused due to the loss of excitatory input from the GC. In the BICC model, classical conditioning can be achieved without the PU but not without the IPN, which is consistent with Lavond and Steinmetz (1989).

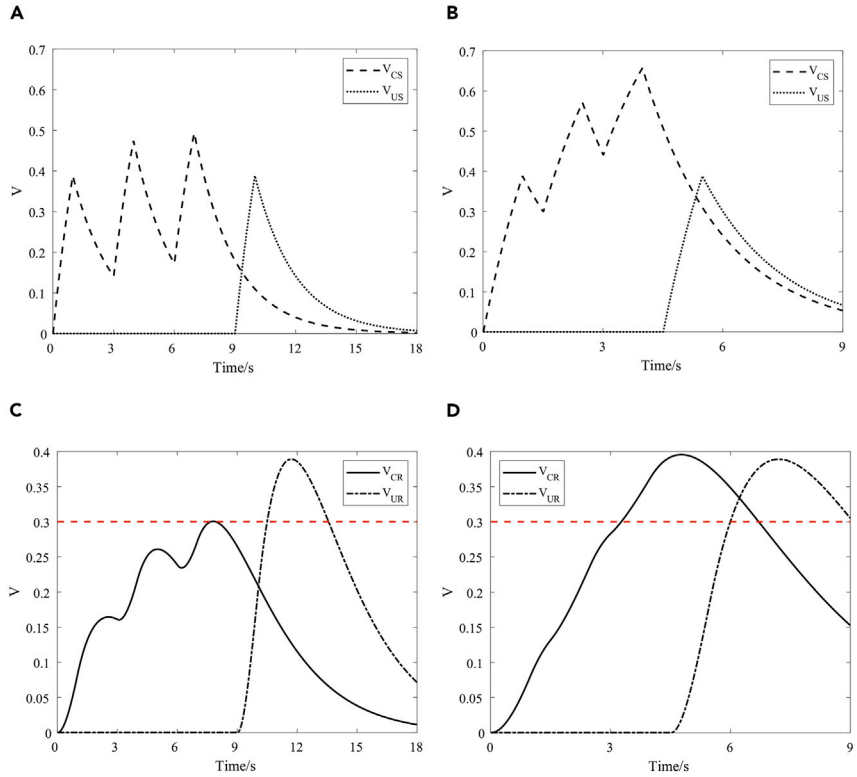


Figure 7. The result of speed generalization experiment in real environment

(A–D) The V_{CR} is the firing rate of response neuron when it receives CS. For easy comparison, we show the firing rate of response neuron when it receives only US (V_{UR}). The red line indicates the response threshold. Compared with (A and B), with the increase of speed, the density of CS increased, so the V_{CS} is stronger. (C) The firing rate of CR and UR at 50% speed. It is easy to see that after training, the robot could perform CR before UR. (D) The firing rate of CR and UR at 100% speed. It is easy to see that without training, the robot could perform CR before UR.

In this article, we propose a BICC model and use 11 classical conditioning experiments to validate this model. The results of experimental validations in a simulation environment and on a humanoid robot indicate that this model can handle almost all classical conditioning experiments and endows the robot with the ability to establish a *CR*.

Limitations of the study

Our model cannot reproduce the experiment of latent inhibition (Lubow and Moore, 1959), spontaneous recovery (Pavlov, 1927), and unblocking (Bradfield and Mcnally, 2008). The latent inhibition effect is that a familiar stimulus takes longer to build an association to response than a new stimulus. Our model cannot reproduce the latent inhibition experiment because the model doesn't distinguish between familiar stimulus and new stimulus. The spontaneous recovery effect is the reappearance of a response that had been extinguished. Our model cannot reproduce this experiment, and we think that the spontaneous recovery effect may require involvement of more brain regions. The unblocking effect is that the responding of the blocked CS_2 increases by increase in the intensity or duration of the *US*, or increase in the number of the *US*. Our model failed in the unblocing effect experiment, if the CS_1 has built an association to response, the CS_2 cannot build the association, no matter how the *US* changes. In the future, we will improve our model to reproduce more experiments.

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Yi Zeng (yi.zeng@ia.ac.cn).

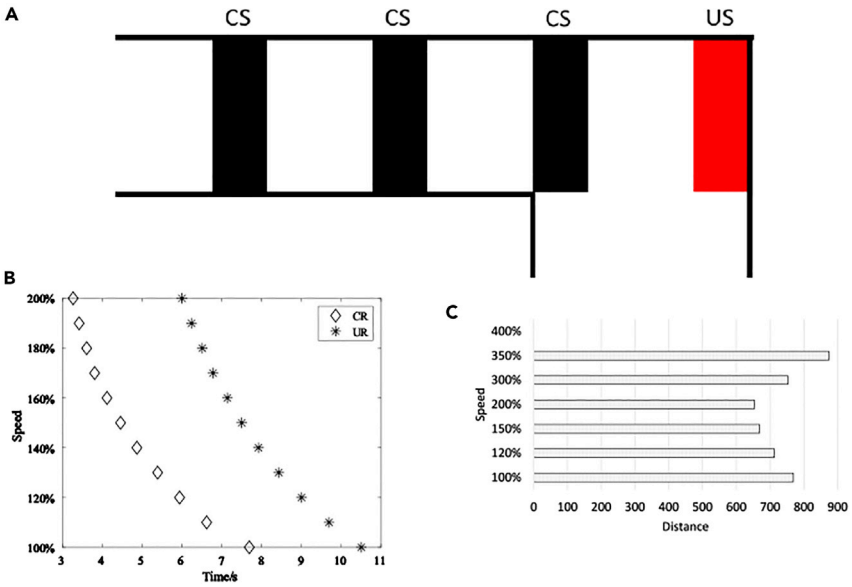


Figure 8. Speed generalization experiment in simulation environment

(A) Simulation environment. The black square is the CS, and the red square is the US. The US indicates that the robot should turn right to avoid leaving the runway.

(B) The induced time of CR and UR at different speeds. It easy to see that the robot could perform CR before UR.

(C) The induced distance of CR at different speeds. The robot successfully passed the experiment with a maximum speed of 350%. When the speed reached 400%, the CS could not induce the CR, and the experiment failed.

Materials availability

This study did not generate new unique reagents.

Data and code availability

The MATLAB scripts can be downloaded from the GitHub repository of the Brain-Inspired Cognitive Engine at Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences: https://github.com/Brain-Inspired-Cognitive-Engine/BICC.

METHODS

All methods can be found in the accompanying Transparent methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101980.

ACKNOWLEDGMENTS

This work is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB32070100), the new generation of artificial intelligence major project of the Ministry of Science and Technology of the People's Republic of China (Grant No. 2020AAA0104305), the Beijing Municipal Commission of Science and Technology (Grant No. Z181100001518006), the Key Research Program of Frontier Sciences, CAS (Grant No. ZDBS-LY-JSC013), the CETC Joint Fund (Grant No. 6141B08010103), and the Beijing Academy of Artificial Intelligence (BAAI).

AUTHOR CONTRIBUTIONS

Conceptualization, Y. Zhao and Y. Zeng; Methodology, Y. Zhao, Y. Zeng, and G.Q.; Software, Y. Zhao and G.Q.; Validation, Y. Zhao and Y. Zeng; Formal Analysis, Y. Zhao and Y. Zeng; Investigation, Y. Zhao and Y. Zeng; Data Curation, Y. Zhao; Writing – Original Draft, Y. Zhao and Y. Zeng; Writing – Review & Editing

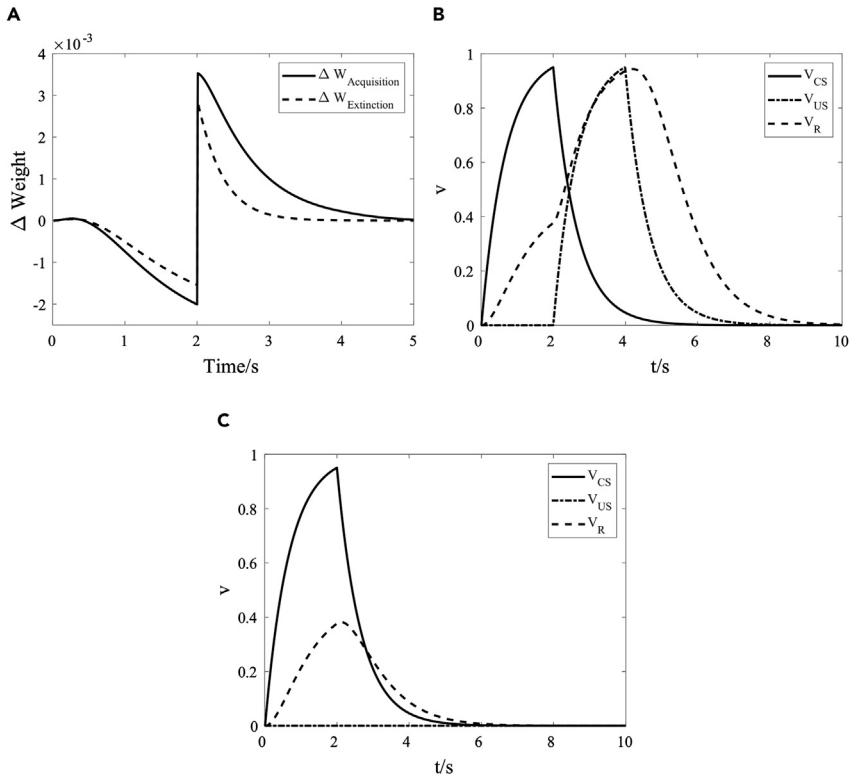


Figure 9. The synaptic weight changing and firing rates of neurons in acquisition and extinction experiments

(A) Single synaptic weight changes in a single trial in the acquisition and extinction experiments.

(B) Firing rates of neurons in the acquisition experiment. The *CS* is presented at 0 *s* and ends at 2 *s*. The *US* is presented at 2 *s* and ends at 4 *s*.

(C) The *CS* is presented at 0 *s* and ends at 2 *s* and there is no *US*.

Draft, Y. Zhao and Y. Zeng; Visualization, Y. Zhao; Supervision, Y. Zeng; Project Administration, Y. Zeng; Funding Acquisition, Y. Zeng.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 12, 2020
Revised: November 24, 2020
Accepted: December 16, 2020
Published: January 22, 2021

REFERENCES

Angulo, R., Bustamante, J., Estades, V., Ramirez, V., and Jorquera, B. (2020). Sex differences in cue competition effects with a conditioned taste aversion preparation. Front. Behav. Neurosci. *14*, 107.

Antonietti, A., Casellato, C., D'Angelo, E., and Pedrocchi, A. (2017). Model-driven analysis of eyeblink classical conditioning reveals the underlying structure of cerebellar plasticity and neuronal activity. IEEE Trans. Neural Netw. Learn. Syst. *28*, 2748–2762.

Balkenius, C., and Moren, J. (1998). Computational models of classical conditioning: a comparative study. From Anim. Animats *5*, 348–353.

Balkenius, C., and Morén, J. (1999). Dynamics of a classical conditioning model. Auton. Robots *7*, 41–56.

Bengio, Y., Mesnard, T., Fischer, A., Zhang, S., and Wu, Y. (2017). Stdp-compatible approximation of backpropagation in an energy-based model. Neural Comput. *29*, 555–577.

Bi, G., and Poo, M. (2001). 'Synaptic modification by correlated activity: Hebb's postulate revisited'. Annu. Rev. Neurosci. *24*, 139–166.

Bradfield, L., and Mcnally, G.P. (2008). Unblocking in pavlovian fear conditioning. J. Exp. Psychol. Anim. Behav. Process. *34*, 256.

Hansel, C., Linden, D.J., and D'Angelo, E. (2001). Beyond parallel fiber ltd: the diversity of synaptic and non-synaptic plasticity in the cerebellum. Nat. Neurosci. *4*, 467–475.

Harry Klopf, A. (1988). A neuronal model of classical conditioning. Psychobiology *16*, 85–125.

Herreros, I., Maffei, G., Brandi, S., Sánchez-Fibla, M., and Verschure, P.F.M.J. (2013). Speed generalization capabilities of a cerebellar model on a rapid navigation task. In 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IEEE), pp. 363–368.

Hogri, R., Bamford, S.A., Taub, A.H., Magal, A., Del Giudice, P., and Mintz, M. (2015). A neuro-inspired model-based closed-loop neuroprosthesis for the substitution of a cerebellar learning function in anesthetized rats. Sci. Rep. *5*, 8451.

Johansson, C., and Lansner, A. (2002). An Associative Neural Network Model of Classical Conditioning (Dept. Numerical Analysis and Computing Science, KTH).

Kleim, J.A., Freeman, J.H.,J., Bruneau, R., Nolan, B.C., Cooper, N.R., Zook, A., and Walters, D. (2002). Synapse formation is associated with memory storage in the cerebellum. Proc. Natl. Acad. Sci. U S A *99*, 13228–13231.

Koekkoek, S.K., Hulscher, H.C., Dortland, B.R., Hensbroek, R.A., Elgersma, Y., Ruigrok, T.J., and De Zeeuw, C.I. (2003). Cerebellar ltd and learning-dependent timing of conditioned eyelid responses. Science *301*, 1736–1739.

Lavond, D.G., and Steinmetz, J.E. (1989). Acquisition of classical conditioning without cerebellar cortex. Behav. Brain Res. *33*, 113–164.

Liu, S., and Ding, Y. (2008). An adaptive network policy management framework based on classical conditioning. In 2008 7th World Congress on Intelligent Control and Automation (IEEE), pp. 3336–3340.

Liu, J., Lu, Y., and Chen, J. (2008). A conditional reflex model based bayesian network (crmbbn). In 2008 Fourth International Conference on Natural Computation (IEEE), pp. 3–8.

Lubow, R.E., and Moore, A.U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. J. Comp. Physiol. Psychol. *52*, 415–419.

Matzel, L.D., Schachtman, T.R., and Miller, R.R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. Learn. Motiv. *16*, 398–412.

Pavlov, I.P. (1927). Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex (Oxford University Press).

PinenO, O., Urushihara, K., and Miller, R.R. (2005). Spontaneous recovery from forward and backward blocking. J. Exp. Psychol. Anim. Behav. Process. *31*, 172.

Pugh, J.R., and Raman, I.M. (2006). Potentiation of mossy fiber epscs in the cerebellar nuclei by nmda receptor activation followed by postinhibitory rebound current. Neuron *51*, 113–123.

Pugh, J.R., and Raman, I.M. (2008). Mechanisms of potentiation of mossy fiber epscs in the cerebellar nuclei by coincident synaptic excitation and inhibition. J. Neurosci. *28*, 10549–10560.

Qiao, G., Du, H., and Zeng, Y. (2017). A quaternionic rate-based synaptic learning rule derived from spike-timing dependent plasticity. In Advances in Neural Networks - ISNN 2017 (Springer International Publishing), pp. 457–465.

Rescorla, R.A., and Wagner, A.R. (1972). A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Classical Conditioning: Current Research and Theory, *2* (Appleton-Century-Crofts), pp. 64–99.

Schmajuk, N.A., and DiCarlo, J.J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. Psychol. Rev. *99*, 268.

Schneiderman, N., Fuentes, I., and Gormezano, I. (1962). Acquisition and extinction of the classically conditioned eyelid response in the albino rabbit. Science *136*, 650–652.

Schneiderman, N., and Gormezano, I. (1964). Conditioning of the nictitating membrane of the rabbit as a function of cs-us interval. J. Comp. Physiol. Psychol. *57*, 188.

Smith, M.C., Coleman, S.R., and Gormezano, I. (1969). 'Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward cs-us intervals'. J. Comp. Physiol. Psychol. *69*, 226–231.

Steuber, V., Mittmann, W., Hoebeek, F.E., Silver, R.A., De Zeeuw, C.I., Hausser, M., and De Schutter, E. (2007). Cerebellar ltd and pattern recognition by purkinje cells. Neuron *54*, 121–136.

Sutton, R.S. (1990). Time-derivative models of pavlovian reinforcement. Learning and computational neuroscience: Foundations of adaptive, *vol. 6* (MIT Press), pp. 497–537.

Sutton, R.S., and Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. Psychol. Rev. *88*, 135.

Sutton, R.S., and Barto, A.G. (1987). A temporal-difference model of classical conditioning. In Proceedings of the Ninth Annual Conference of the Cognitive Science Society (Lawerence Erlbaum), pp. 355–378.

Takehara-Nishiuchi, K. (2018). The anatomy and physiology of eyeblink classical conditioning. Curr. Top. Behav. Neurosci. *37*, 297–323.

Ten Brinke, M.M., Boele, H.J., and De Zeeuw, C.I. (2019). Conditioned climbing fiber responses in cerebellar cortex and nuclei. Neurosci. Lett. *688*, 26–36.

Wang, D., Smith-Bell, C.A., Burhans, L.B., O'Dell, D.E., Bell, R.W., and Schreurs, B.G. (2018). Changes in membrane properties of rat deep cerebellar nuclear projection neurons during acquisition of eyeblink conditioning. Proc. Natl. Acad. Sci. U S A *115*, E9419–E9428.

Weeks, A.C., Connor, S., Hinchcliff, R., LeBoutillier, J.C., Thompson, R.F., and Petit, T.L. (2007). Eye-blink conditioning is associated with changes in synaptic ultrastructure in the rabbit interpositus nuclei. Learn. Mem. *14*, 385–389.

Wetmore, D.Z., Jirenhed, D.A., Rasmussen, A., Johansson, F., Schnitzer, M.J., and Hesslow, G. (2014). Bidirectional plasticity of purkinje cells matches temporal features of learning. J. Neurosci. *34*, 1731–1737.

Zhang, W., and Linden, D.J. (2006). Long-term depression at the mossy fiber-deep cerebellar nucleus synapse. J. Neurosci. *26*, 6935–6944.

Zuo, G., Yang, B., and Ruan, X. (2005). The cognitive behaviors of a spiking-neuron based classical conditioning model. In Advances in Intelligent Computing (Springer Berlin Heidelberg), pp. 939–948.

iScience, Volume *24*

Supplemental Information

Brain-inspired classical conditioning model

Yuxuan Zhao, Yi Zeng, and Guang Qiao

Supplementary Materials

This PDF file includes:

Transparent Methods
Supplementary References
Supplementary Figures S1-S4
Supplementary Tables S1

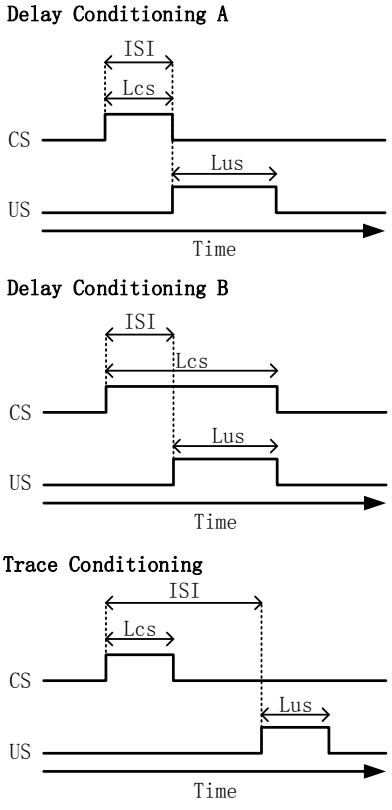
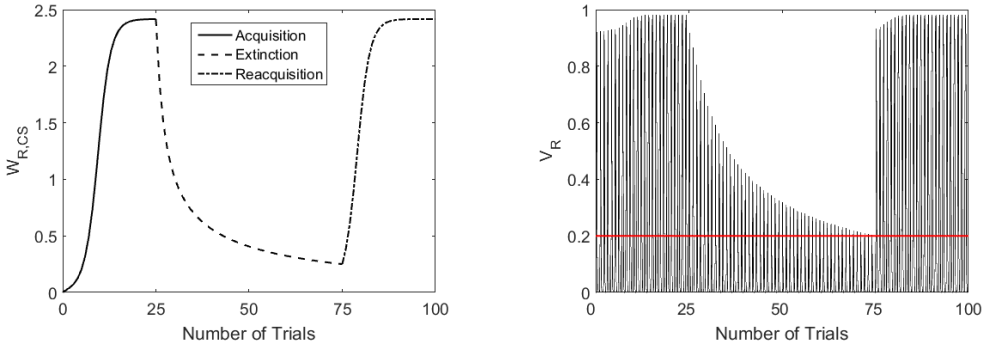


Figure S1: Three Inter-stimulus Interval conditionings. **(A)** In delay conditioning A, the inter-stimulus interval (ISI) is equal to the length of the CS (L_{CS}). **(B)** In delay conditioning B, the inter-stimulus interval (ISI) is equal to the difference between the length of the CS (LCS) and the length of the US (L_{US}). **(C)** In trace conditioning, the inter-stimulus interval (ISI) is equal to the difference between the onset of US and CS. Related to Figure 2A.



A B

Figure S2: Total Experiment. The first 25 trials is acquirement experiment, the subsequent 50 trials is extinction experiment, and the last 25 trials is reacquisition experiment. The red line in figure response is the response threshold which we setted as 0.2. **(A)** The synaptic weight changing between NP_{CS} and NP_R. **(B)** The firing rate of Response neuron populations in total experiment, and the red line is the threshold of response. Related to Figure 2c and Figure 2d.

Transparent Methods

Computational Model

The architecture of the BICC model is shown in Figure 1.

The change in the firing rate of a neuron that has received a conditioned stimulus (CS) or an unconditioned stimulus (US) in this model can be described as shown in Equation 1. $S_j(t)$ represents stimulus j $\in \{CS_1, CS_2, ..., \mathrm{CS}_N, \mathrm{US}\}$ at time t, and $\mathrm{V}_j(\mathrm{t})$ represents the firing rate of the neuron. We set $S_j(t)$ to 1 if the stimulus is present at time t and set it to 0 if the stimulus has ended. The firing rate of the neuron will increase when the stimulus is present and decay when the stimulus has ended. C is the parameter that controls the rate of increase and decrease.

$$\Delta V_j(t) = -C \times (V_j(t) - S_j(t)) \tag{1}$$

Synaptic plasticity occurs in two areas: between *GC (granule cell)* and *PU (Purkinje cell)* and between the *PN (pontine nuclei)* and *IPN (interpositus nucleus)*. The Purkinje cell is the sole output of the cerebellar cortex and inhibits the neurons in the IPN (interpositus nucleus) (Takehara-Nishiuchi, 2018), and it shows spontaneous firing due to its intrinsic membrane properties, as supported by evidence from several studies (Häusser and Clark, 1997; Raman and Bean, 1997). In the cerebellar cortex, the firing rate of the Purkinje cell depends on its own spontaneous firing, the stimulation from the CS and US, and the inhibitory information from inhibitory interneurons, as described in Equation 2.

$$V_{pu}(t+1) = V_{pu}(t) + \Delta V_{pu}(t) \tag{2}$$

where

$$\Delta V_{pu}(t) = -C \times (V_{pu}(t) - tanh(V_{pu-self}(t) + \sum_k (W_{gc_k}(t) \times V_{gc_k}(t) + W_{int.n} \times V_{int.n_k}(t)) + W_{io} \times V_{io}(t)))$$

$V_{pu}(t)$ is the firing rate of the Purkinje cell at time t, and $V_{pu-self}$ is the spontaneous firing rate of the Purkinje cell itself. For k $\in \{CS_1, CS_2, ..., \mathrm{CS}_N\}$, $W_{gc_k}(t)$ is the synaptic weight between the GC (granule cell) and the PU (Purkinje cell), V_{gc_k} is the firing rate of the GC (granule cell), $W_{int.n}$ is the synaptic weight between the Int.N (inhibitory interneuron) and the PU (Purkinje cell), and $V_{int.n_k}$ is the firing rate of the Int.N (inhibitory interneuron). W_{io} is the synaptic weight between the IO (inferior olive) and the PU (Purkinje cell), and $V_{io}(t)$ is the firing rate of the IO (inferior olive). $W_{int.n}$ and W_{io} remain unchanged in this model.

In the IPN (interpositus nucleus), the firing rate of the neuron depends on the inhibitory information from the Purkinje cell and the stimulation from the CS and US, as described in Equation 3.

$$V_{ipn}(t+1) = V_{ipn}(t) + \Delta V_{ipn}(t) \tag{3}$$

where

$$\Delta V_{ipn}(t) = -C \times (V_{ipn}(t) - tanh(W_{pu} \times V_{pu}(t) + \sum_k W_{pn_k}(t) \times V_{pn_k}(t) + W_{io} \times V_{io}(t)))$$

$V_{ipn}(t)$ is the firing rate of the neuron in the IPN (interpositus nucleus) at time t. W_{pu} is the synaptic weight between the IPN (interpositus nucleus) and the PU (Purkinje cell), and $V_{pu}(t)$ is the firing rate of the Purkinje cell. $W_{pn_k}(t)$ is the synaptic weight between the PN (pontine nuclei) and the IPN (interpositus nucleus), and $V_{pn_k}(t)$ is the firing rate of the PN (pontine nuclei). W_{io} is the synaptic weight between the IO (inferior olive) and the IPN (interpositus nucleus), and $V_{io}(t)$ is the firing rate of IO (inferior olive). W_{pu} and W_{io} remain unchanged in this model.

The synaptic plasticity in this model is defined as shown in Equation 4. W_{ij} in this model represents the synaptic weight between postsynaptic neuron i and presynaptic neuron population j. The postsynaptic neurons contain the neurons in the PU (Purkinje cell) and the IPN (interpositus nucleus), and the presynaptic neurons contain the neurons in the GC (granule cell) and the PN (pontine nuclei).

$$W_{ij}(T+1) = W_{ij}(T) + \Delta W_{ij}(T) \tag{4}$$

where

$$\Delta W_{ij}(T) = K_{ij}(T) \int_T^{T+1} \Delta w_{ij}(t) dt$$

$K_{ij}(T)$ in this model represents the number of synapses between postsynaptic neuron i and presynaptic neuron population j. $\Delta w_{ij}(t)$ is the incremental weight of a single synapse, which is calculated using Equation 5.

$$\Delta w_{ij}(t) = \alpha V_i(t)V_j(t) + \beta V_i'(t)V_j(t) + \gamma V_i(t)V_j'(t) \tag{5}$$

where

$$\alpha = \int_{-\infty}^{+\infty} f(u)du, \beta = \int_{-\infty}^{0} uf(u)du, \gamma = -\int_{0}^{+\infty} uf(u)du$$

$f(u)$ is the STDP function reported in (Bi and Poo, 2001; Song et al., 2000; Shouval et al., 2002). And the general shape of the STDP function is shown as Figure S3.

Equation 5 is a quaternionic-rate-based synaptic learning rule derived from spike-timing-dependent plasticity, which we proposed in (Qiao et al., 2017). It is easy to see that there are three main terms that affect the changes in synaptic weight, namely, the product of the pre- and postsynaptic neuron firing rates, the product of the presynaptic neuron firing rate and the time derivative of the postsynaptic neuron firing rate, and the product of the postsynaptic neuron firing rate and the time derivative of the presynaptic neuron firing rate, with corresponding coefficients α, β and γ. Equation 5 is used to eliminate the incompatibility of spike-timing-dependent plasticity and classical conditioning on a temporal scale, because the effective time window in spike-timing-dependent plasticity is on the millisecond scale, such as 40 ms, but it can be as long as several seconds in classical conditioning.

In the learning process of a biological system, changes occur not only in the form of synaptic weight changes but also in the forms of synaptic growth and synaptic elimination (Johnston et al., 2009; Bailey et al., 2004). The change in the number of synapses, $\Delta K_{ij}(T+1)$, is defined in Equation 6.

$$\Delta K_{ij}(T+1) = \int_{T}^{T+1} k \times \Delta w_{ij}(t)dt \times W_{ij}(T) \tag{6}$$

If $W_{ij}(T) > 0$, this means that most of the synapses between neuron population j and neuron population i are excitatory connections. If $\Delta w_{ij}(T) > 0$, this means that the strength of excitatory connections increases and the number of excitatory synapses simultaneously increases, resulting in synaptic growth. If $\Delta w_{ij}(T) < 0$, this means that the strength of excitatory connections decreases and the number of excitatory synapses simultaneously decreases, resulting in synaptic elimination.

If $W_{ij}(T) < 0$, this means that most of the synapses between neuron population j and neuron population i are inhibitory connections. If $\Delta w_{ij}(T) > 0$, this means that the strength of inhibitory connections decreases and the number of inhibitory synapses simultaneously decreases, resulting in synaptic elimination. If $\Delta w_{ij}(T) < 0$, this means that the strength of inhibitory connections increases and the number of inhibitory synapses simultaneously increases, resulting in synaptic growth.

Parameter Setting

We make the stimulus continues for 2s, and the characteristic behaviors of a neuron for which we set the parameter C to 1/200, 2/200, 3/200, 4/200 in this model is shown in Figure S4A. When the stimulus ends, the slow decrease of the neuron's firing rate is conducive to the establish the conditioned response in trace condition, but if it takes too long for the neuron's firing rate to decay to zero, it is not conducive to the convergence of the model, such as the parameter is 1/200. The neuron's characteristic behaviors are similar when the parameter is 2/200, 3/200, 4/200, and each parameter could make the model establish the association between the stimulus and the response. So we set the parameter C to 3/200 in this model which makes the firing rate of the neuron increase rapidly and decrease slowly.

The parameters α, β and γ calculated by the STDP function that reported in (Bi and Poo, 2001) can't make the model reproduce the classical experiments. So, we calculate these parameters using the incremental weights in three states.

We convert the Equation 1 to a function with t as the variable, and the converted function is shown in Equation 7. The t_c is the continues time of the stimulus.

$$V_j(t) = \begin{cases} 1-(1-C)^t & stimulus \quad presented \\ (1-C)^{t-t_c} \times (1-(1-C)^{t_c}) & stimulus \quad ended \end{cases} \tag{7}$$

We simplified the equation for calculating the firing rate of postsynaptic neuron i, as shown in Equation 8.

$$V_i(t+1) = V_i(t) - C \times (V_i(t) - tanh(W \times V_j(t))) \tag{8}$$

We calculate the parameters α, β and γ using the incremental weights in three states, they are the acquisition state, the acquisition stable state and the extinction state. In the acquisition state and the acquisition stable state, the CS is presented at $0s$ and ends at $2s$, and the US is presented at $2s$ and ends at $4s$. In the extinction state, only a CS is presented at $0s$ and ends at $2s$, without a US.

Because of the non-linear transformation in Equation 8, we calculate the firing rate of the postsynaptic neuron i in three states respectively, and then use the fourth-order Fourier function to fit the firing rate curve. So, we can get a function of postsynaptic neuron's firing rate with t as the variable, as shown in Equation 9. And the parameters in three states are shown in Table S1.

$$\begin{aligned} V_i(t) = a0 + a1 \times cos(t \times w) + b1 \times sin(t \times w) + a2 \times cos(2 \times t \times w) + b2 \times sin(2 \times t \times w) \\ +a3 \times cos(3 \times t \times w) + b3 \times sin(3 \times t \times w) + a4 \times cos(4 \times t \times w) + b4 \times sin(4 \times t \times w) \end{aligned} \tag{9}$$

The $\Delta W_{ij}(T)$ need to be greater than zero, equal to zero, and less than zero in the acquisition state, the acquisition stable state and the extinction state respectively. With the Equation 5 and set the $K_{ij}(T)$ to 1, we can get a three-variable linear equation set as shown in Equation 10.

$$\begin{cases} 38.1403 \times \alpha + 0.1898 \times \beta - 0.5674 \times \gamma > 0 & acquisition\ \ state \\ 147.1433 \times \alpha + 0.5810 \times \beta - 0.5676 \times \gamma = 0 & acquisition\ \ stable\ \ state \\ 143.7313 \times \alpha + 0.4563 \times \beta - 0.4424 \times \gamma < 0 & extinction\ \ state \end{cases} \tag{10}$$

We set the $\Delta W_{ij}(T)$ to 0.25, 0, -0.1 in the acquisition state, the acquisition stable state and the extinction state respectively. By solving the equations, we can get the $\alpha = -0.0035$, $\beta = 0.3441$ and $\gamma = -0.5510$. Consider that the fourth-order Fourier function cannot fully fit the firing rate curve of the postsynaptic neuron, we select parameters of $\alpha = -0.0035$, $\beta = 0.35$, $\gamma = -0.55$ in this model. And the STDP function calculated according to these parameters is shown in Figure S4b.

When the $W_{ij}(T)$ and $\Delta w_{ij}(T)$ are constant, the k determines the rate of synaptic growth and elimination. The larger the k, the faster the synaptic growth and extinction in the process of acquisition and extinction, and the model needs less trials to complete the acquisition or extinction. The smaller k is, the slower the rate of synaptic growth and extinction in the process of acquisition and extinction, and more trials are required for the model to complete acquisition or extinction. We select parameters of $k = 8$, so that the model can complete acquisition or regression with a relatively appropriate number of trials (such as 25), and facilitate observation of experimental phenomena.

Bibliography

Bailey, C. H., Kandel, E. R. and Si, K. (2004), 'The persistence of long-term memory: a molecular approach to self-sustaining changes in learning-induced synaptic growth', *Neuron* **44**(1), 49–57.

Bi, G. and Poo, M. (2001), 'Synaptic modification by correlated activity: Hebb's postulate revisited', *Annual Review of Neuroscience* **24**, 139–166.

ChrolCannon, Gruning and Jin (2012), The emergence of polychronous groups under varying input patterns, plasticity rules and network connectivities, *in* 'Neural Networks (IJCNN), The 2012 International Joint Conference on'.

Häusser, M. and Clark, B. A. (1997), 'Tonic synaptic inhibition modulates neuronal output pattern and spatiotemporal synaptic integration', *Neuron* **19**(3), 665–678.

Johnston, M. V., Ishida, A., Ishida, W. N., Matsushita, H. B., Nishimura, A. and Tsuji, M. (2009), 'Plasticity and injury in the developing brain', *Brain and Development* **31**(1), 1–10.

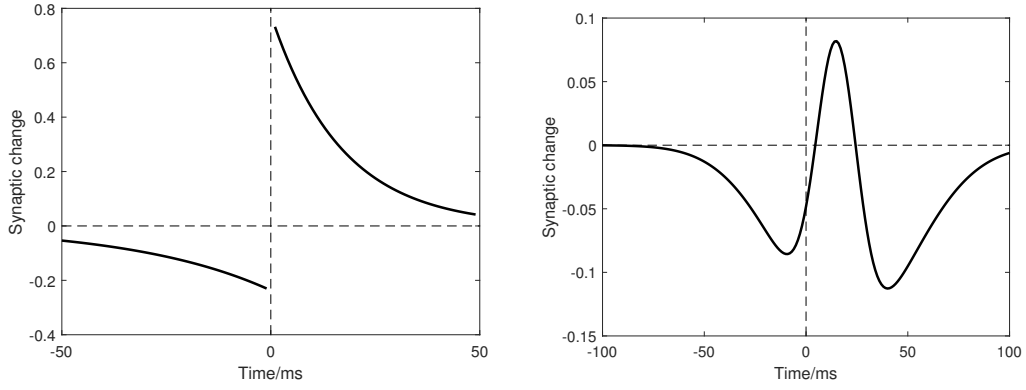
Qiao, G., Du, H. and Zeng, Y. (2017), A quaternionic rate-based synaptic learning rule derived from spike-timing dependent plasticity, *in* 'Advances in Neural Networks - ISNN 2017', Springer International Publishing, pp. 457–465.

Raman, I. M. and Bean, B. P. (1997), 'Resurgent sodium current and action potential formation in dissociated cerebellar purkinje neurons', *The Journal of Neuroscience* **17**(12), 4517–4526.

Shouval, H. Z., Bear, M. F. and Cooper, L. N. (2002), 'A unified model of nmda receptor-dependent bidirectional synaptic plasticity', *Proceedings of the National Academy of Sciences of the United States of America* **99**(16), 10831–10836.

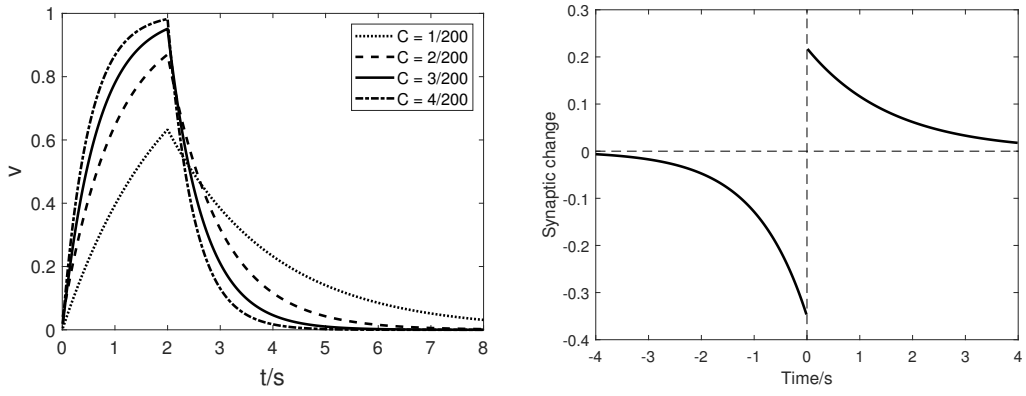
Song, S., Miller, K. D. and Abbott, L. F. (2000), 'Competitive hebbian learning through spike-timing-dependent synaptic plasticity', *Nature Neuroscience* **3**(9), 919–926.

Takehara-Nishiuchi, K. (2018), 'The anatomy and physiology of eyeblink classical conditioning', *Current Topics in Behavioral Neurosciences* **37**, 297–323.



A B

Figure S3: The general shape of the STDP function. The horizontal axis is the spike time difference between the postsynaptic spike and the presynaptic spike. **(A)** The biphasic STDP. If $\Delta t > 0$, $\Delta w = A^+ exp(-\Delta t/\tau^+)$. If $\Delta t < 0$, $\Delta w = A^- exp(-\Delta t/\tau^-)$. As described in (Bi and Poo, 2001), the $A^+ = 0.777$, $A^- = -0.237$, $\tau^+ = 16.8$, $\tau^- = -33.7$. **(B)** The triphasic STDP. $\Delta w = A^+ exp(-(\Delta t - 15)^2/\tau^+) - A^- exp(-(\Delta t - 20)^2/\tau^-)$. As described in (ChrolCannon et al., 2012), the $A^+ = 0.23$, $A^- = 0.15$, $\tau^+ = 200$, $\tau^- = 2000$. Related to Figure 1.



A B

Figure S4: The characteristic behaviors of a neuron and the STDP function in our model. **(A)** The characteristic behaviors of a neuron that the parameter C is 1/200, 2/200, 3/200, 4/200. **(B)** The STDP function in our model. The parameters in this function are $A^+ = 0.2183$, $A^- = -0.35$, $\tau^+ = 1.5873$, $\tau^- = -1$. Related to Figure 1.

Table S1: The parameters in three states. Related to Figure 1.

	a0	a1	b1	a2	b2	a3	b3	a4	b4	w
acquisition	0.4668	-0.4386	-0.0180	0.0378	0.0354	-0.0111	-0.0052	-0.0252	-0.0098	0.0082
acquisition stable	0.1939	-0.2142	0.7971	-0.0813	0.2019	0.0773	0.0761	-0.0302	0.0227	0.0049
extinction	0.0136	-0.1265	0.7026	-0.0780	0.4613	0.0729	0.0784	0.0494	-0.0269	0.0052