Title:

Counteracting memories acquired in parallel by limited pairings in eyeblink conditioning

Authors:

Ryota Iwase[1], Shin-ya Kawaguchi[1, *]

Affiliations:

[1] Department of Biophysics, Graduate School of Science, Kyoto University; Oiwake-cho, Kitashirakawa, Sakyo-ku, Kyoto, 606-8502, Japan

*Corresponding author. E-mail: kawaguchi@nb.biophys.kyoto-u.ac.jp

Abstract

Animals adapt to environments by learning even when experience is limited, which is critical for their survival in nature. However, the relationship between experience richness and learning efficiency remains elusive. Here we investigated this issue using mice eyeblink conditioning with three distinct training paradigms. Surprisingly, 1/9th of sensory pairings than conventional ones were not only sufficient for acquisition of conditioned responses, relying on the activity of deep cerebellar nuclei as usual, but also resulted in more efficient learning. On the other hand, the learning established by limited training was more susceptible for extinction. Modeling of establishment and extinction processes, coupled with experimental validation of the model prediction, unveiled that memories for execution of eyeblink and its suppression are independently formed irrespective of their temporal order.

Introduction

Animals adapt to changing environments by learning from experience, whose richness has been related to the efficiency of adaptive behavior (Grant & Schipper, 1952; Thomas & Wagner, 1964). However, it is unclear how animals behave with severely limited exposure to reinforcement experience. Here we studied this issue using classical eyeblink conditioning in mice, which has been extensively studied for a century to assess learning ability and the underlying neural circuits (Cason, 1922; Christian & Thompson, 2003; Heiney et al., 2014). Repeated presentation of conditioned stimulus (CS; e.g., tone or light) coupled with eyelid reflex-inducing unconditioned stimulus (US; e.g., air puff to cornea) typically at high probability (~90%) makes animals learn to blink in response to the initially neutral CS. Such acquired responses through training are called conditioned responses (CRs). Although previous studies demonstrated that even the halved pairing of sensory inputs is effective to establish CRs (Thomas & Wagner, 1964; Leonard & Theios, 1967; Gormezano & Coleman, 1975; Buchanan et al., 1997; Kehoe et al., 2008), it remains elusive whether much rare experience is still effective for the CR acquisition. In this study, by changing the enrichment of sensory pairings, we quantitatively investigated the acquisition and extinction of associative eyeblink learning focusing on the probability and behavior dynamics, and found that much less and/or low density of associative sensory experience is not just enough but rather more effective for the establishment of well-timed eyelid closure, relying on the identical area of cerebellar nuclei. On the other hand, excessive training enhanced the resistance of CR memory to extinction. Further, detailed analysis of experimental data yielded a simple model to explain the acquisition and extinction processes for learned behavior through parallel formation of memories for the execution of eyeblink and its suppression.

Results

Eyeblink conditioning with limited paired sensory stimuli

To examine whether limited experience of paired sensory stimuli can establish eyeblink conditioning in mice, we first changed the ratio of two sensory stimuli: tone (290 ms) and air puff (to a cornea, 40 ms) as CS and US, respectively (Fig. 1A, B). Mice were head-fixed on a treadmill during experiments (Fig. 1C) and their eyelid movements were recorded with a monochrome camera, and the eyelid closure level was quantitated by calculating the area size of an eyeball (Fig. 1D; Heiney et al., 2014; for details see Materials & Methods). Air puff delivered via a glass tube or a needle elicited quick eyeblink responses (unconditioned responses, URs), whereas tone did not affect the eyelid (Fig. 1D, before). For association of two sensory events in mice, combined and co-terminating sensory inputs of CS and US in time series are typically provided with high ratio: 90 paired & 10 CS-alone trials ($90\%_{CS\text{-}US}$, Fig. 1A). In line with previous studies (Heiney et al., 2014; Fiocchi et al., 2022), CRs were evident in ~10 out of 100 trials on average even on the 1st day, gradually becoming frequent (~ 50%) and larger by ~ day4, which remained nearly constant till day10 (Fig. 1E, F), although showing variable time courses in individual mice (Fig. S1A). On the other hand, surprisingly, even when the pairing of CS and US was lowered to 10% (i.e., 10 paired & 90 CS-alone per day: $10\%_{CS\text{-}US}$, Fig. 1A), mice acquired CRs during 10 days training (Fig. 1E, middle) although hitting a plateau around 30% on day6 (Fig. 1F). Thus, even 9 times less sensory inputs succeeded in the establishment of eyeblink conditioning, while the final probability for the CR triggering was apparently lower with delay in reaching equilibrium (Thomas & Wagner, 1964; Leonard & Theios, 1967; Gormezano & Coleman, 1975; Buchanan et al., 1997; Kehoe et al., 2008).

Two factors could underly the lower percentage of CRs (%CR) in the $10\%_{CS\text{-}US}$ group: insufficient paired inputs or excessive presentation of unpaired tone. To discriminate these,

another paradigm '11 trials' training was examined (10 paired & 1 CS-alone per day), where the limited paired sensory experience was the same as the $10\%_{CS\text{-}US}$, but the excessive tone-alone stimulation was omitted (Fig. 1A, B). As shown in Fig. 1E, F and S1A, even the daily 10 CS-US pairings effectively established CRs at ~50% in 6 days training, similarly to the daily 100 trials in the $90\%_{CS\text{-}US}$ (Fig. 1F), although slight delay in reaching plateau. Success rate of learning in individual mice in three groups was similar, although slightly low in the $10\%_{CS\text{-}US}$ group (Fig. S1B). These results altogether indicate that small number of paired sensory stimuli (10 pairings per day) are sufficient for the CR acquisition, while additional massive tone-alone events only lowers final probability of learned behavior. Notably, considering that the probability of air puff was just 10% out of total tone presentations in the $10\%_{CS\text{-}US}$, the final %CR as ~30% might be rather regarded as exaggerated occurrence of acquired behavior.

When the averaged %CR was plotted as a function of total CS-US pairings, amazingly, the $10\%_{CS\text{-}US}$ and the 11 trials groups exhibited very quick establishment of CRs compared to the conventional $90\%_{CS\text{-}US}$ group (Fig. 1G). The courses of increase in %CR were fitted by the following equation:

$$\%CR = P_{max} * \frac{1}{1 + (\frac{L_{50}}{N_{paired}})^n}$$

, where P_{max}, L_{50}, n, and N_{paired} represent the maximal %CR, the number of paired stimuli for the half-maximal learning, the steepness of learning, and the number of given paired trials. In sharp contrast to the similar n in three groups, L_{50} reflecting the number of sensory inputs needed for learning was ~6 folds higher in the $90\%_{CS\text{-}US}$ group, suggesting less efficiency of the conventional training in mice unlike in rabbits (Thomas & Wagner, 1964; Leonard & Theios, 1967; Gormezano & Coleman, 1975). Step-by-step pattern of %CR increases were evident in the $90\%_{CS\text{-}US}$ group during initial 4 days, implying some daily upper limit for learning such as a

gradual attenuation of its efficiency during training (Tait et al., 1983; Kehoe & White, 2002). The %CR in the $10\%_{CS\text{-}US}$ group was compatible initially in each day to the $90\%_{CS\text{-}US}$ and the 11 trials groups, followed by gradual decline during a daily session (Fig. 1G, inset). Thus, tone-alone trials might not deteriorate the acquisition, but distract the expression of learned CRs.

The acquired eyeblink responses were similarly sophisticated in timing and amplitude although the averaged eyelid closures in the $10\%_{CS\text{-}US}$ group were apparently smaller because of larger variability of peak timing (Fig. 2A-C). The CR amplitude grew up gradually during days of training and showed faster saturation in the $10\%_{CS\text{-}US}$ and the 11 trials groups in relation to the number of paired trials, while onset/peak timings got quickly optimized in all groups (Fig. S2A-E) (Balsam et al., 2002; Ohyama & Mauk, 2001; Kehoe et al., 2008; Kehoe et al., 2014; Heiney et al., 2014; Fiocchi, et al., 2022). With a closer look, the $10\%_{CS\text{-}US}$ group occasionally exhibited rapid and complete eyelid closure (RCEC), reaching 98% eyelid closure during 80-180 ms from the tone onset (Fig. S3A, B). Interestingly, the initial dynamics of RCEC looked similar to that of spontaneous eye blinking (Fig. S3C). As a result of more RCEC occurrence, well-timed blinking relative to the timing of air puff, which was regarded to be cerebellum-dependent CRs, was slightly rarer in the $10\%_{CS\text{-}US}$ group (Fig. S3D), while total probability of eyeblink responses including CRs and RCEC was almost compatible in three groups (Fig. S3E).

Critical role of DCN in CR expression

Deep cerebellar nuclei (DCN) plays an important role in expression of CRs in rabbits (McCormick & Thompson, 1984; Yeo et al., 1985; Christian & Thompson, 2003) and mice (Sakamoto & Endo, 2008; Heiney et al., 2014). To examine whether the CRs acquired with $10\%_{CS\text{-}US}$ also rely on DCN, we pharmacologically inhibited neuronal activity by infusing muscimol, a selective $GABA_A$ receptor agonist, to the anterior interpositus part of DCN through a cannula

implanted in the already CR-acquired mice (Fig. 3A). As shown in Fig. 3B, cannula position and estimated area for diffusion of muscimol were similar in the two groups (90% and 10% CS-US pairing groups). Muscimol administration (evaluated by at least 30 trials given before and after its infusion) potently abolished the CR expression in both the 90% and 10% CS-US pairing groups, followed by almost complete recovery on the next day (Fig. 3C). Averaged traces for eyelid closure indicated almost complete elimination of CRs by muscimol, but not by artificial cerebrospinal fluid (aCSF), without affecting the URs upon air puff (Fig 3D-F). Taken these results together, it was revealed that CR is expressed depending on the DCN activity irrespective of the altered enrichment of paired sensory experience.

Scatter plots of averaged eyelid closure amplitude against %CR exhibited a tendency of the $90\%_{CS\text{-}US}$ showing positive relation between the amplitude and the percentage of CRs (Fig. S4A, left), which was preserved even after the DCN inactivation. In contrast, the $10\%_{CS\text{-}US}$ group showed less correlation between %CR and eyelid closure amplitude, sometimes exhibiting unaltered amplitude even after the DCN inactivation (Fig. S4A, right). Furthermore, RCEC sometimes evident in the $10\%_{CS\text{-}US}$ group tended to resist to the DCN inactivation (Fig. S4B). Thus, limited experience and disturbance by lots of CS-alone inputs in the $10\%_{CS\text{-}US}$ group might be partly compensated by recruiting neural circuits other than the cerebellar circuit for eliciting eyelid closure (Kaminer et al., 2011; Kaminer et al., 2015; Bologna et al., 2024).

Distinct extinction dynamics of CRs acquired with altered enrichment of sensory experience

In order to evaluate the qualitative identity of CR memories acquired by different patterns of experience, we examined memory-extinction paradigm. Following 10 days of acquisition training, 100 tone-alone (i.e. 100 CS-alone) trials were daily applied for 4 days. CR occurrence gradually decreased during the extinction session with clearly different kinetics depending on the pattern of

acquisition (Fig. 4A-C). As evident in Fig. 4B, looking at gradual changes during a session, %CR substantially recovered at the beginning of individual daily sessions, known as spontaneous recovery (Robleto et al., 2004; Delamater, 2004; Bouton 2004; Rescorla, 2004; Thanellou & Green, 2011). The CR extinction profiles (Fig. 4C) could be fitted with single exponential curves, as characterized by the following equation:

$$\%CR = P_{ext} * \exp\left(-\frac{N_{CS}}{\tau_{ext}}\right)$$

, where P_{ext}, τ_{ext} and N_{CS} represent the initial value of %CR, the factor for resistance of CR to be erased, and the number of CS-alone trials during each day, respectively. The CRs established by rich paired stimuli seem to exhibit ~4 times greater resistance to extinction in mice, as in rabbits (Thomas & Wagner, 1964; Gormezano & Coleman, 1975), but distinct from humans in which CRs acquired with partial reinforcement showed higher resistance (Humphreys, 1939; Grant & Hake, 1951; Grant & Schipper, 1952; Moore & Gormezano, 1963). In clear contrast to the %CR attenuation, eyelid closure amplitude was kept high during the extinction session in all groups although the 11 trials group showed slight decline (Fig. 4D, E). Furthermore, interestingly, eye blinking like spontaneous responses tended to be frequently observed during extinction (Fig. S3C, 4F, G). As for the timing control, response onset showed delay during extinction although peak latency exhibited large variabilities among individuals (Fig. 4H), as previously reported (Medina et al., 2002; Kehoe et al., 2014).

Modeling CR acquisition and extinction through integrating total history of sensory inputs

To shed light on the underlying mechanisms of distinct dynamics of CR acquisition and extinction with altered sensory enrichment, we attempted to develop a model that explains our observations (see Figs. 1 and 4). Several lines of evidences in our experiments such as the apparently preserved eyeblink dynamics and the daily spontaneous recovery during extinction (see Fig. 4), suggest that

the memory encoding CR execution and that to suppress it seem to be separately established, as previously suggested (Robleto et al., 2004; Delamater, 2004; Bouton, 2004; Rescorla, 2004). Thus, we built a model assuming parallel formation of counteracting memories for CRs, as follows (here, in modeling section, we denote %CR as CR_{ratio}):

$$CR_{ratio} = v1_{2nd} + v1_{1st} * (1 - g * v2)$$

$$v1 = v1_{1st} + v1_{2nd}$$

, where $v1$ is the memory for CR execution consisting of $v1_{1st}$ and $v1_{2nd}$ with distinct susceptivity to the CR suppressing memory, $v2$. g represents a gate for $v2$, increasing as a function of total amount of CS inputs (i.e. total trial number) (Fig. 5A; for details see Materials and Methods). The larger $v1$ results in higher CR_{ratio}, while the larger $v2$ leads to smaller CR_{ratio}. $v1_{1st}$, $v1_{2nd}$ and $v2$ were assumed to build up trial by trial based on the Rescorla and Wagner model (Rescorla & Wagner, 1972) with some modification, defined as follows:

$$\Delta v1_{1st} = \alpha 1_{1st} * \beta * (\lambda 1 - v1) - \alpha 1_{2nd} * v1_{1st} \qquad \text{[upon CS-US]}$$

$$\Delta v1_{2nd} = \alpha 1_{2nd} * v1_{1st} \qquad \text{[upon CS-US]}$$

$$\Delta v2 = \alpha 2 * (\lambda 2 - v2) \qquad \text{[upon CS-alone]}$$

, where $\alpha 1_{1st}$, $\alpha 1_{2nd}$ and $\alpha 2$ represent rate constants for formation of memories $v1_{1st}$, $v1_{2nd}$ and $v2$, respectively. $v1_{1st}$ was assumed to be formed upon the CS-US pairings at a rate $\alpha 1_{1st}$ multiplied by an efficiency factor β, and to be gradually turned into the extinction-resistant $v1_{2nd}$ at a rate $\alpha 1_{2nd}$. Stronger resistance of CR to extinction in the $90\%_{\text{CS-US}}$ was made possible by this scheme (see Fig. 4C). Daily upper limit of learning, inferred from Fig. 1G, was implemented in the model by gradual decline of the $\alpha 1_{1st}$ efficiency factor β as total paired trials increase within a session (Fig. S5A).

Both the acquisition and extinction processes of CR_{ratio} were simulated in the model, and distinct dynamics for three experimentally-tested patterns were nicely reproduced (Fig. 5B-

F). The $90\%_{CS\text{-}US}$ group hit daily plateau of CR acquisition, whereas the 11 trials group showed more efficient learning as observed in experiments. In the $10\%_{CS\text{-}US}$ group, even after reaching apparent saturation, CR_{ratio} exhibited daily quick increase which disappeared subsequently because of gradual gate of slowly summating *v2* to counteract $v1_{1st}$ via activation of *g* by repeated trials (Fig. 5F).

We also conceived another model in which the suppressing memory, *v2* was short-lived and decayed by the next session, rather than long-lasting (Fig. 5E, S6B). As shown in Fig. S6A, this model also succeeded in reproducing experimental observations. In this version, notably, daily spontaneous recovery during extinction sessions can be explained just by the complete refreshment of suppressing *v2* memory irrespective of the inclusion of a gating mechanism of *v2*. However, the simulated decline of CR_{ratio} following quick increase each day in the $10\%_{CS\text{-}US}$ was strongly affected by the gating mechanism: eliminating the gate made the daily CR_{ratio} decline more rapid than actual experimental data (Fig. S6C). Thus, in any cases, the suppressing memory *v2* should be adequately controlled by the online experience of sensory inputs, as has been suggested in a previous study showing that the extinction session speeded down the decline of %CR with occasional paired sensory inputs (Gibbs et al., 1978).

To evaluate the qualitative difference of learning dynamics generated by two distinct *v2,* long- or short-lasting one, we attempted to obtain a prediction from two distinct models. Repetitive CS-alone trials given for days in prior strongly suppressed the later CR_{ratio} increase in the model with long-lasting *v2*, but not in the model assuming short-lasting *v2* (Fig. 5G). This prediction was tested by behavior experiments. As shown in Fig. 5G and H, the 100 CS-alone trials for 4 days prior to the start of 90% CS-US pairing paradigm (the Pre group) potently inhibited the initial growth of CRs (Lubow & Moore, 1959; Schmajuk et al., 1994; Nicholson & Freeman, 2002) with similar eyelid closure dynamics to the other experience patterns (Fig. S7A).

The increase of CR ratio showed quite similar pattern to that in the $10\%_{CS\text{-}US}$ groups rather than the $90\%_{CS\text{-}US}$, characterized by daily decline (Fig. 5I). The zig-zag pattern of CR_{ratio} dynamics was more evident in the model with long-lasting *v2* (Fig. 5G). Thus, even before the acquisition of CRs, amazingly, the prior experience of CS-alone inputs seem to exert negative effect not only on the later apparent acquisition but also on the high-probability execution of CRs after acquisition, in a similar manner to the CR memory established by the limited pairings. Collectively, we conclude that the association of CS and US forms a memory to execute CRs, while CS-alone experience forms a counteracting memory to suppress CRs in parallel and independently from the situation of the target positive memory.

Discussion

Eyeblink conditioning is widely conserved in mammals, showing specie-dependent various relations between the training pattern and the learning dynamics. Humans show %CR closely related to the ratio of pairings (Grant & Hake, 1951; Grant & Schipper, 1952; Ross, 1959; Moore & Gormezano, 1963), while rabbits supra-linearly exhibit CRs upon training with smaller possibility of sensory pairings (Thomas & Wagner, 1964; Leonard & Theios, 1967; Gormezano & Coleman, 1975; Buchanan et al., 1997). As demonstrated here, even much smaller 10% pairing is sufficient for the CR acquisition in mice (see Fig. 1F), showing rather inversed relationship between the enrichment of associative experience and the learning efficiency (see Fig. 1G), which contrasts to the previous findings in rabbits typically showing equal or less efficiency with partial reinforcement (Thomas & Wagner, 1964; Leonard & Theios, 1967; Gormezano & Coleman, 1975). Apparent slower growth of %CR in mice than in rabbits during days of trainings (Thomas & Wagner, 1964; Leonard & Theios, 1967; Gormezano & Coleman, 1975; Heiney et al., 2014;

Fiocchi et al., 2022), implies more potent upper limit of learning during a certain period in mice, spoiling surplus reinforcement experience. This could be responsible for the apparent higher learning efficiency upon limited reinforcement. Indeed, learning efficiency can be estimated higher also in rabbits with smaller density of training, when compared with a training paradigm accompanied with daily plateau or decline of %CR (Tait et al., 1983; Kehoe & White, 2002).

We demonstrated a critical role of cerebellar circuit in expression of learned eyeblink irrespective of the enrichment of reinforcement (Fig. 3; McCormick et al., 1982; Yeo et al., 1985; Krupa et al., 1993; Christian & Thompson, 2003). Long-term depression (LTD) at parallel fiber (PF) synapses on a Purkinje cell (PC) has been focused as a substrate of memory formation for CRs (Koekkoek et al., 2003). In addition, plasticity at mossy fiber (MF) synapses on neurons in DCN also plays a role, contributing to higher resistance to extinction presumably requiring more experience for its establishment compared to PF-PC LTD (Garcia & Mauk, 1998; Medina et al., 2001; Ohyama et al., 2006; Antonietti et al., 2016; Antonietti et al., 2017; Broersen et al., 2023). Given these findings, the PF-PC LTD and plasticity at MF-DCN synapses might correspond to the extinction-sensitive ($v1_{1st}$) and -resistant ($v1_{2nd}$) memories for CRs in our hypothetical model (see Fig. 5A). In line with previous studies suggesting an independent memory formation for extinction of learned behavior (Robleto et al., 2004; Delamater, 2004; Bouton, 2004; Rescorla, 2004), our experiments and modeling together highlighted an importance of the memory counteracting CR execution, not only for the extinction, but also for the fine-tuned probability to execute learned eyelid closure. In addition, surprisingly, our data suggest that the counteracting memory can be somehow long-lasting and formed even in prior to the formation of positive memory to execute eyeblink, the latter of which has been regarded as the latent inhibition to date (Lubow & Moore, 1959; Schmajuk et al., 1994; Nicholson & Freeman, 2002). The neuronal mechanism and location of such memory trace are unclear at present, but previous elegant studies

showing a critical role of cerebello-olivary circuits (Medina et al., 2002; Kim et al., 2020) and hippocampus (Schmaltz & Theios, 1972; Akase et al., 1989) in CR extinction might provide a clue to clarify the key element to encode the enrichment of sensory inputs into the integrated fine-tuned memory for motor control.

Materials & Methods

Animal usages

Adult male FVB/NJcl mice, ≥ 8 weeks old, were used in all experiments (n = 86). Mice were housed singly with *ad libtum* access to food and water in a room with 12 hours of light/dark cycle. All animal procedures were performed by the same experimenter in accordance with guidelines on animal experimentation in Kyoto University and approved by the local committee in Graduate School of Science, Kyoto University.

Surgery

Custom-made aluminum head-plates were mounted on mice before experiments. Mice were anesthetized with intraperitoneal injection (0.2-0.3 ml/mouse) of mixture of 0.9% ketamine (Daiichisankyo) and 0.2% xylazine (Bayer). The scalp was cut off to expose the skull and mice were stereotaxically fixed on an operating table (Narishige). The underlying membranous tissue was removed with cotton swabs and small holes were made around bregma using a drill. Screws were loosely tightened to the holes. The head plate was placed on the skull surface and fixed to the screws and the skull with dental cement (UNIFAST Trad, GC). Mice had at least 30 hours of recovery period before experiments started.

To pharmacologically inactivate DCN, a guide cannula (RWD Life Science) was

implanted for some mice (n = 12) together with a head plate (see Fig. 3A). A small hole (~6.0 mm posterior, ~2.0 mm lateral from bregma; Paxinos & Flanklin, 2008) was drilled and the cannula tip was approached to the brain surface using a stereotaxic micromanipulator. The cannula was inserted 1.8-2.3 mm vertically and fixed with dental cement. An internal dummy cannula was inserted to the guide cannula to prevent clogging.

Eyeblink conditioning, acquisition and extinction

The basic apparatus was composed of a self-made treadmill on which mice were head-fixed during experiments, a camera (CS135MU, Thorlabs) for eyelid closure recording, a speaker for tone stimuli, a solenoid for air stimuli, and a PC (XPS8940, Dell) for recording and stimuli control (Fig. S8A; Heiney et al., 2014). The treadmill was composed of cylindrical rotating rods and rubber sheet. The implanted head-plate was fixed to a manipulator (SM-19, Narishige) ~2 cm above the treadmill with a screw (Fig. 1C). The camera was mounted on a metal rod to monitor a left eye. Mice were illuminated by white LED light and eyelid movements were recorded with 100 frames/sec. The speaker was placed back left side of mice. A gas cylinder was connected to the solenoid and air was delivered to a left cornea when the solenoid opened. The camera, speaker and solenoid were controlled in timing with Python through an AD/DA converter (DIO-8/8B-UBT, Y2 Corporation).

Mice were randomly divided into three groups with distinct trainings ($90\%_{CS\text{-}US}$: n = 22; $10\%_{CS\text{-}US}$: n = 20; 11 trials: n = 22), as follows (Fig. 1A): $90\%_{CS\text{-}US}$ group, 10 blocks of 10 trials per day, with each block composed of 9 paired & 1 CS-alone trials; $10\%_{CS\text{-}US}$ group, 10 blocks of 1 paired & 9 CS-alone trials; 11 trials group, 10 paired & 1 CS-alone trials per day. Each mouse had at least 6 days of training. Mice had habituation period without recording and stimuli delivery for ≥ 15minutes on the treadmill before training each day. The CS (tone, 3.5 kHz, 75 dB, 290

ms) was co-terminated with the US (air puff, 15-30 psi, 40 ms) in CS-US paired trials, on the other hand, the tone CS was given without the US presentation in CS-alone trials (Fig. 1A). The US was delivered via a glass tube (i.d. = 0.5 mm) or a 21 gauge needle placed ~2 mm away from a mouse cornea (Fig. 1C) and given in almost identical manner with gradual increase in pressure as days of training proceeded (Fig. S8B). The order of paired and CS-only trials was randomly determined in each block. The inter-trial interval was set to 10-24 sec randomly for each trial.

Some mice ($90\%_{CS\text{-}US}$, n = 8; $10\%_{CS\text{-}US}$, n = 9; 11 trials, n = 7) had 4 consecutive days of extinction paradigm, 100 CS-alone trials/day, following the 10 consecutive days of acquisition. Other mice (n = 8) had exposure to daily 100 CS-alone trials for 4 consecutive days prior to the $90\%_{CS\text{-}US}$ trainings (Pre group). When training was accidentally interrupted in the first block because of system' s error, although very rare, daily traning was started over again.

Data analysis

Recorded videos for eyelid movements were converted into numerical trace by calculating the area of an eyeball in individual frames (Fig. 1D; Heiney et al., 2014). Full eyelid open/close corresponded to 0/100% respectively. Reflection of light and infiltration of whisker were approximately calibrated. Trials were omitted from analysis when significant activities other than eyelid closures were observed, or recording or saving of eyelid movements was not performed properly. Eyelid closure amplitude was obtained by subtraction of pre-CS baseline from the maximal eyelid closure during 80-250 ms after the CS onset (for CS-US paired trials) or 80-470 ms (for CS-alone trials). When the amplitude was $\geq$ 15%, this trial was regarded as one with a response and further analysis was conducted. Onset of eyelid closure was visually identified. The α-startle, a startle reflex to tone, was carefully removed from the analysis by visual identification. The responses with averaged eyelid closure amplitude during 220-250 ms $\geq$ 15% and larger than

that during 120-150 ms, were regarded as CRs.

Pharmacological inactivation of DCN neurons

DCN neurons in CR-learned mice (90%$_{CS-US}$: n = 6; 10%$_{CS-US}$: n = 6) were inactivated by muscimol (1 or 5 mM, Nakalai Tesque) administration through an implanted guide cannula (see Surgery). Muscimol was diluted in aCSF (NaCl, 125 mM, $NaHCO_3$, 25 mM, NaH_2PO_4, 1.25 mM, KCl, 2.5 mM, D-glucose, 10 mM, $CaCl_2 \cdot 2H_2O$, 2 mM, $MgCl_2 \cdot 6H_2O$, 2 mM, adjusted to pH ~7.4 and ~300 mOsm). An internal cannula connected to a microsyringe (701N 10 μl, Hamilton) through silicon tube was inserted into the implanted guide cannula and muscimol (0.5-5 μl) was infused (~1 μl/min). The internal cannula was removed carefully and the examination of eyelid movements resumed within 15 mins. Mice were imposed ≥ 30 trials before/after the infusion. In some mice (90%$_{CS-US}$: n = 3; 10%$_{CS-US}$: n = 3), aCSF was infused without muscimol on another day for negative control. The session was omitted from analysis when UR or ordinal behavior significantly changed after the infusion.

Muscimol diffused area and the cannula position were confirmed in individual mice. Acriflavine, a fluorescent dye, (0.5-5 μl, 0.01%, Wako), was infused in the same way as the muscimol infusion (90%$_{CS-US}$: n = 5; 10%$_{CS-US}$: n = 6) after pharmacological inactivation experiments. Mice were euthanized by cervical dislocation and the head plates were detached from the skulls. Whole brains were carefully removed and fixed in 4% paraformaldehyde/PBS (paraformaldehyde: Nakalai Tesque; PBS: Takara Bio) overnight and soaked for >12 hours in ~20% sucrose for cryoprotection. Horizontal sections were prepared at ~50 μm thickness using a cryostat (SM2000R, Leica). These sections were immediately mounted (VECTARSHIELD Mounting Medium, Funakoshi) and imaged with a confocal fluorescent microscope (FV1000, Olympus). The area of muscimol diffusion was estimated from the acriflavine fluorescence. The

position of small holes in the sections was regarded as the positions of cannula.

Modeling

Learning model for CR acquisition and extinction was constructed, and the dynamics of probability of CR occurrence was simulated under the Python programming language (Figs. 5, S5, S6). Memories for CR execution (*v1*) and suppression (*v2*) were assumed to be formed independently. As for *v1*, $v1_{1st}$ and $v1_{2nd}$ with distinct sensitivity to CR extinction or suppression were assumed based on the fact that the CRs showed resistance to extinction depending on the enrichment of paired sensory experiences (see Fig. 4C), together with the previous studies suggesting two candidate synapses encoding the CR-triggering memory with distinct resistance for extinction: synapses on cerebellar PCs and those in DCN (Garcia & Mauk, 1998; Medina et al., 2001; Ohyama et al., 2006; Antonietti et al., 2016; Antonietti et al., 2017; Broersen et al., 2023). On the other hand, trials-dependent gradual suppression by the memory *v2*, was implemented by the gate control '*g*', which made the spontaneous recovery possible in the model.

Taken together, the CR_{ratio} was expressed as a function of three memories, as follows:

$$CR_{ratio} = v1_{2nd} + v1_{1st} * (1 - g * v2)$$

$$v1 = v1_{1st} + v1_{2nd}$$

Three memories, $v1_{1st}$, $v1_{2nd}$ and *v2*, changed trial by trial based on Rescorla-Wagner model (Rescorla & Wagner, 1972), as follows:

$$\Delta v1_{1st} = \alpha 1_{1st} * \beta * (\lambda 1 - v1) - \alpha 1_{2nd} * v1_{1st} \qquad \text{[upon CS-US]}$$

$$\Delta v1_{2nd} = \alpha 1_{2nd} * v1_{1st} \qquad \text{[upon CS-US]}$$

$$\Delta v2 = \alpha 2 * (\lambda 2 - v2) \qquad \text{[upon CS-alone]}$$

, where $\lambda 1$ and $\lambda 2$ are saturation levels for *v1* and *v2*, and set to 1 here. $\alpha 1_{1st}$, $\alpha 1_{2nd}$, and $\alpha 2$ are rate constants for formation of $v1_{1st}$, $v1_{2nd}$ and *v2*, respectively. β was assumed here to make

the CS-US pair-dependent formation of $v1_{1st}$ sensitive to the total number of daily paired trials within a session (N_{paired}) (Fig. S5A):

$$\beta = \frac{1}{1 + (\frac{N_{paired}}{\tau_\beta})^{n_\beta}}$$

, where τ_β and n_β define the steepness of experience-dependent slowing of $v1_{1st}$ formation.

Gating of suppressing memory *v2, g*, was also implemented as follows:

$$g = 1 - \frac{g_0}{1 + (\frac{N_{trial}}{\tau_{gate}})^{n_{gate}}}$$

, where N_{trial} is the total trial number in a session. g_0 is a factor to decide the initial gate active at the beginning of each session. τ_{gate} and n_{gate} define the steepness of the trials-dependent gate opening, and the former was defined to change depending on experience as follows (Fig. S5B):

$$\tau_{gate} = A_{gate} + B_{gate} * \frac{1}{1 + \left(\frac{p_\tau}{p(pair)_{exp}}\right)^{n_\tau}}$$

, where A_{gate} and B_{gate} are the initial value and the experience-dependent maximal increment of τ_{gate}, respectively. p_τ and n_τ are constants to define the steepness of τ_{gate} change. Expected paring ratio, $p(pair)_{exp}$, was assumed to be updated trial by trial as follows (Fig. S5C):

$$\Delta p(pair)_{exp} = \gamma * \left(1 - p(pair)_{exp}\right) \qquad \text{[upon CS-US]}$$

$$\Delta p(pair)_{exp} = \gamma * \left(0 - p(pair)_{exp}\right) \qquad \text{[upon CS-alone]}$$

, where γ is the rate constant for change of $p(pair)_{exp}$.

Individual memories were assumed to fade over time (t) as defined by the following equations:

$$\Delta v1_{1st} = -(1 - \frac{1}{1 + \left(\frac{t}{\varepsilon_{v1_1st}}\right)^{n_{decay}}}) * v1_{1st} + (1 - \frac{1}{1 + \left(\frac{t}{\varepsilon_{v1_2nd}}\right)^{n_{decay}}}) * v1_{2nd}$$

$$\Delta v1_{2nd} = -(1 - \frac{1}{1 + \left(\frac{t}{\varepsilon_{v1_2nd}}\right)^{n_{decay}}}) * v1_{2nd}$$

$$\Delta v2 = -(1 - \frac{1}{1 + \left(\frac{t}{\varepsilon_{v2}}\right)^{n_{decay}}}) * v2$$

, where ε_{v1_1st}, ε_{v1_2nd}, ε_{v2}, and n_{decay} are time constants of memory decay and their steepness. Daily decay of each was calculated before each session started in simulation using 1440 min, corresponding to 24 hours, for time t. Other variables, β, g, and $p(pair)_{exp}$ are refreshed to initial values at the beginning of each day (see Supplementary Table 1).

Simulation was performed trial by trial, giving CS-US paired or CS-alone trials as in the experiments. The 2 models assuming long- or short- lasting *v2* suppressing memory, and a model without gate control were performed by changing parameters (see Supplementary Table 1).

Statistics

Data was presented as mean ± SEM unless otherwise stated. Paired t-test, Tukey-Kramer test and two-way ANOVA were performed with Microsoft Excel or Rstudio (Posit) for statistical analysis. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$ were regarded as significant. Approximate curves were obtained with Igor Pro8 (WaveMetrics) using Chi-square method.

References

1. Akase, E., Alkon, D. L., & Disterhoft, J. F. (1989). Hippocampal lesions impair memory of short-delay conditioned eye blink in rabbits. *Behavioral Neuroscience*, *103*(5), 935–943. https://doi.org/10.1037/0735-7044.103.5.935

2. Antonietti, A., Casellato, C., D'Angelo, E., & Pedrocchi, A. (2017). Model-Driven Analysis of Eyeblink Classical Conditioning Reveals the Underlying Structure of Cerebellar Plasticity and Neuronal Activity. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(11), 2748–2762. https://doi.org/10.1109/TNNLS.2016.2598190
3. Antonietti, A., Casellato, C., Garrido, J. A., Luque, N. R., Naveros, F., Ros, E., Dangelo, E., & Pedrocchi, A. (2016). Spiking Neural Network With Distributed Plasticity Reproduces Cerebellar Learning in Eye Blink Conditioning Paradigms. *IEEE Transactions on Biomedical Engineering*, *63*(1), 210–219. https://doi.org/10.1109/TBME.2015.2485301
4. Balsam, P. D., Drew, M. R., & Yang, C. (2002). Timing at the Start of Associative Learning. *Learning and Motivation*, *33*(1), 141–155. https://doi.org/10.1006/lmot.2001.1104
5. Bologna, M., Paparella, G., Valls-Solé, J., Hallett, M., & Berardelli, A. (2024). Neural control of blinking. *Clinical Neurophysiology*, *161*, 59–68. https://doi.org/10.1016/j.clinph.2024.02.023
6. Bouton, M. E. (2004). Context and Behavioral Processes in Extinction. *Learning & Memory, 11(5),* 485-494. http://www.learnmem.org/cgi/doi/10.1101/ lm.78804
7. Broersen, R., Albergaria, C., Carulli, D., Carey, M. R., Canto, C. B., & De Zeeuw, C. I. (2023). Synaptic mechanisms for associative learning in the cerebellar nuclei. *Nature Communications*, *14*, 7459. https://doi.org/10.1038/s41467-023-43227-w
8. Buchanan, S. L., Penney, J., Tebbutt, D., & Powell, D. A. (1997). Lesions of the mediodorsal nucleus of the thalamus and classical eyeblink conditioning under less-than-optimal stimulus conditions: Role of partial reinforcement and interstimulus interval. *Behavioral Neuroscience*, *111*(5), 1075–1085. https://doi.org/10.1037/0735-7044.111.5.1075
9. Cason, H. (1922). The conditioned eyelid reaction. *Journal of Experimental Psychology*, *5*(3), 153–196. https://doi.org/10.1037/h0074822
10. Christian, K. M., & Thompson, R. F. (2003). Neural Substrates of Eyeblink Conditioning: Acquisition and Retention. *Learning & Memory*, *10*(6), 427–455. https://doi.org/10.1101/lm.59603
11. Delamater, A. (2004). Experimental extinction in Pavlovian conditioning: Behavioural and neuroscience perspectives. *The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*, *57*, 97–132. https://doi.org/10.1080/02724990344000097
12. Fiocchi, F. R., Dijkhuizen, S., Koekkoek, S. K. E., De Zeeuw, C. I., & Boele, H. J. (2022). Stimulus Generalization in Mice during Pavlovian Eyeblink Conditioning. *Eneuro*, *9*(2), ENEURO.0400-21.2022. https://doi.org/10.1523/ENEURO.0400-21.2022
13. Garcia, K. S., & Mauk, M. D. (1998). Pharmacological analysis of cerebellar contributions

to the timing and expression of conditioned eyelid responses. *Neuropharmacology*, *37*(4–5), 471–480. https://doi.org/10.1016/S0028-3908(98)00055-0

14. Gibbs, C. M., Latham, S. B., & Gormezano, I. (1978). Classical conditioning of the rabbit nictitating membrane response: Effects of reinforcement schedule on response maintenance and resistance to extinction. *Animal Learning & Behavior*, *6*(2), 209–215. https://doi.org/10.3758/BF03209603
15. Gormezano, I., & Coleman, S. R. (1975). Effects of partial reinforcement on conditioning, conditional probabilities, asymptotic performance, and extinction of the rabbit's nictitating membrane response. *The Pavlovian Journal of Biological Science*, *10*(1), 13–22. https://doi.org/10.1007/BF03000620
16. Grant, D. A., & Hake, H. W. (1951). Dark adaptation and the Humphreys random reinforcement phenomenon in human eyelid conditioning. *Journal of Experimental Psychology*, *42*(6), 417–423. https://doi.org/10.1037/h0057031
17. Grant, D. A., & Schipper, L. M. (1952). The acquisition and extinction of conditioned eyelid responses as a function of the percentage of fixed-ratio random reinforcement. *Journal of Experimental Psychology*, *43*(4), 313–320. https://doi.org/10.1037/h0057186
18. Heiney, S. A., Wohl, M. P., Chettih, S. N., Ruffolo, L. I., & Medina, J. F. (2014). Cerebellar-Dependent Expression of Motor Learning during Eyeblink Conditioning in Head-Fixed Mice. *The Journal of Neuroscience*, *34*(45), 14845–14853. https://doi.org/10.1523/JNEUROSCI.2820-14.2014
19. Humphreys, L. G. (1939). The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *Journal of Experimental Psychology*, *25*(2), 141–158. https://doi.org/10.1037/h0058138
20. Kaminer, J., Powers, A. S., Horn, K. G., Hui, C., & Evinger, C. (2011). Characterizing the Spontaneous Blink Generator: An Animal Model. *The Journal of Neuroscience*, *31*(31), 11256–11267. https://doi.org/10.1523/JNEUROSCI.6218-10.2011
21. Kaminer, J., Thakur, P., & Evinger, C. (2015). Effects of subthalamic deep brain stimulation on blink abnormalities of 6-OHDA lesioned rats. *Journal of Neurophysiology*, *113*(9), 3038–3046. https://doi.org/10.1152/jn.01072.2014
22. Kehoe, E. J., Ludvig, E. A., Dudeney, J. E., Neufeld, J., & Sutton, R. S. (2008). Magnitude and timing of nictitating membrane movements during classical conditioning of the rabbit (Oryctolagus cuniculus). *Behavioral Neuroscience*, *122*(2), 471–476. https://doi.org/10.1037/0735-7044.122.2.471
23. Kehoe, E. J., Ludvig, E. A., & Sutton, R. S. (2014). Time course of the rabbit's conditioned nictitating membrane movements during acquisition, extinction, and reacquisition. *Learning & Memory*, *21*(11), 585–590. https://doi.org/10.1101/lm.034504.114

24. Kehoe, E. J., & White, N. E. (2002). Extinction revisited: Similarities between extinction and reductions in US intensity in classical conditioning of the rabbit's nictitating membrane response. *Animal Learning & Behavior*, *30*(2), 96–111. https://doi.org/10.3758/BF03192912
25. Kim, O. A., Ohmae, S., & Medina, J. F. (2020). A cerebello-olivary signal for negative prediction error is sufficient to cause extinction of associative motor learning. *Nature Neuroscience*, *23*(12), 1550–1554. https://doi.org/10.1038/s41593-020-00732-1
26. Koekkoek, S. K. E., Hulscher, H. C., Dortland, B. R., Hensbroek, R. A., Elgersma, Y., Ruigrok, T. J. H., & De Zeeuw, C. I. (2003). Cerebellar LTD and Learning-Dependent Timing of Conditioned Eyelid Responses. *Science*, *301*(5640), 1736–1739. https://doi.org/10.1126/science.1088383
27. Krupa, D. J., Thompson, J. K., & Thompson, R. F. (1993). Localization of a Memory Trace in the Mammalian Brain. *Science*, *260*(5110), 989–991. https://doi.org/10.1126/science.8493536
28. Leonard, D. W., & Theios, J. (1967). Classical eyelid conditioning in rabbits under prolonged single alternation conditions of reinforcement. *Journal of Comparative and Physiological Psychology*, *64*(2), 273–276. https://doi.org/10.1037/h0088043
29. Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: The effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*, *52*(4), 415–419. https://doi.org/10.1037/h0046700
30. McCormick, D. A., Clark, G. A., Lavond, D. G., & Thompson, R. F. (1982). Initial localization of the memory trace for a basic form of learning. *Proceedings of the National Academy of Sciences*, *79*(8), 2731–2735. https://doi.org/10.1073/pnas.79.8.2731
31. McCormick, D. A., & Thompson, R. F. (1984). Cerebellum: Essential Involvement in the Classically Conditioned Eyelid Response. *Science*, *223*(4633), 296–299. https://doi.org/10.1126/science.6701513
32. Medina, J. F., Garcia, K. S., & Mauk, M. D. (2001). A Mechanism for Savings in the Cerebellum. *The Journal of Neuroscience*, *21*(11), 4081–4089. https://doi.org/10.1523/JNEUROSCI.21-11-04081.2001
33. Medina, J. F., Nores, W. L., & Mauk, M. D. (2002). Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. *Nature, 416*, 330-333. Doi: 10.1038/416330a
34. Moore, J. W., & Gormezano, I. (1963). Effects of omitted versus delayed UCS on classical eyelid conditioning under partial reinforcement. *Journal of Experimental Psychology*, *65*(3), 248–257. https://doi.org/10.1037/h0040671

35. Nicholson, D., & Freeman, J. (2002). Medial Dorsal Thalamic Lesions Impair Blocking and Latent Inhibition of the Conditioned Eyeblink Response in Rats. *Behavioral Neuroscience*, *116*, 276–285. https://doi.org/10.1037/0735-7044.116.2.276
36. Ohyama, T., & Mauk, M. D. (2001). Latent Acquisition of Timed Responses in Cerebellar Cortex. *The Journal of Neuroscience*, *21*(2), 682–690. https://doi.org/10.1523/JNEUROSCI.21-02-00682.2001
37. Ohyama, T., Nores, W. L., Medina, J. F., Riusech, F. A., & Mauk, M. D. (2006). Learning-Induced Plasticity in Deep Cerebellar Nucleus. *The Journal of Neuroscience*, *26*(49), 12656–12663. https://doi.org/10.1523/JNEUROSCI.4023-06.2006
38. Paxinos, G., & Franklin, K. B. J. (2008). The Mouse Brain in Stereotaxic Coordinates, Compact, Third Edition: The coronal plates and diagrams (Third Edition). *Academic Press*.
39. Rescorla, R. A. (2004). Spontaneous Recovery. *Learning & Memory*, *11*(5), 501–509. https://doi.org/10.1101/lm.77504
40. Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH & Prokasy WF (Editors), Classical Conditioning II: Current Research and Theory.
41. Robleto, K., Poulos, A. M., & Thompson, R. F. (2004). Brain Mechanisms of Extinction of the Classically Conditioned Eyeblink Response. *Learning & Memory*, *11*(5), 517–524. https://doi.org/10.1101/lm.80004
42. Ross, L. E. (1959). The decremental effects of partial reinforcement during acquisition of the conditioned eyelid response. *Journal of Experimental Psychology*, *57*(2), 74–82. https://doi.org/10.1037/h0046965
43. Sakamoto, T., & Endo, S. (2008). GABAA receptors in deep cerebellar nuclei play important roles in mouse eyeblink conditioning. *Brain Research*, *1230*, 125–137. https://doi.org/10.1016/j.brainres.2008.06.079
44. Schmajuk, N. A., Lam, Y.-W., & Christiansen, B. A. (1994). Latent inhibition of the rat eyeblink response: Effect of hippocampal aspiration lesions. *Physiology & Behavior*, *55*(3), 597–601. https://doi.org/10.1016/0031-9384(94)90122-8
45. Schmaltz, L. W., & Theios, J. (1972). Acquisition and extinction of a classically conditioned response in hippocampectomized rabbits (Oryctolagus cuniculus). *Journal of Comparative and Physiological Psychology*, *79*(2), 328–333. https://doi.org/10.1037/h0032531
46. Tait, R. W., Kehoe, E. J., & Gormezano, I. (1983). Effects of US duration on classical conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(1), 91–101. https://doi.org/10.1037/0097-7403.9.1.91
47. Thanellou, A., & Green, J. T. (2011). Spontaneous Recovery But Not Reinstatement of the

Extinguished Conditioned Eyeblink Response in the Rat. *Behavioral Neuroscience*, *125*(4), 613–625. https://doi.org/10.1037/a0023582

48. Thomas, E., & Wagner, A. R. (1964). PARTIAL REINFORCEMENT OF THE CLASSICALLY CONDITIONED EYELID RESPONSE IN THE RABBIT. *Journal of Comparative and Physiological Psychology*, *58*, 157–158. https://doi.org/10.1037/h0048811
49. Yeo, C. H., Hardiman, M. J., & Glickstein, M. (1985). Classical conditioning of the nictitating membrane response of the rabbit. *Experimental Brain Research*, *60*(1), 87–98. https://doi.org/10.1007/BF00237022

Acknowledgments:

We thank Drs T. Inoshita and H. Hirai for the critical reading of the manuscript and helpful comments.

Funding:

This work was supported by Japan Society for Promotion of Science, KAKENHI grants (22H02721, 22K19360, 25H02611, and 25K02362 to S.K.), Takeda Science Foundation (to S.K.), JST SPRING (JPMJSP2110 to R.I.).

Author contributions:

R.I. and S.K. designed the study; R.I. performed the experiments; R.I. analyzed and R.I. and S.K. interpreted the data; all authors wrote and edited the manuscript and approved the final version, and agree to be accountable for that all aspects of the work are appropriately investigated and resolved; all persons designated as authors qualify for authorship, and all those who qualify for authorship are listed.

Competing interests:

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

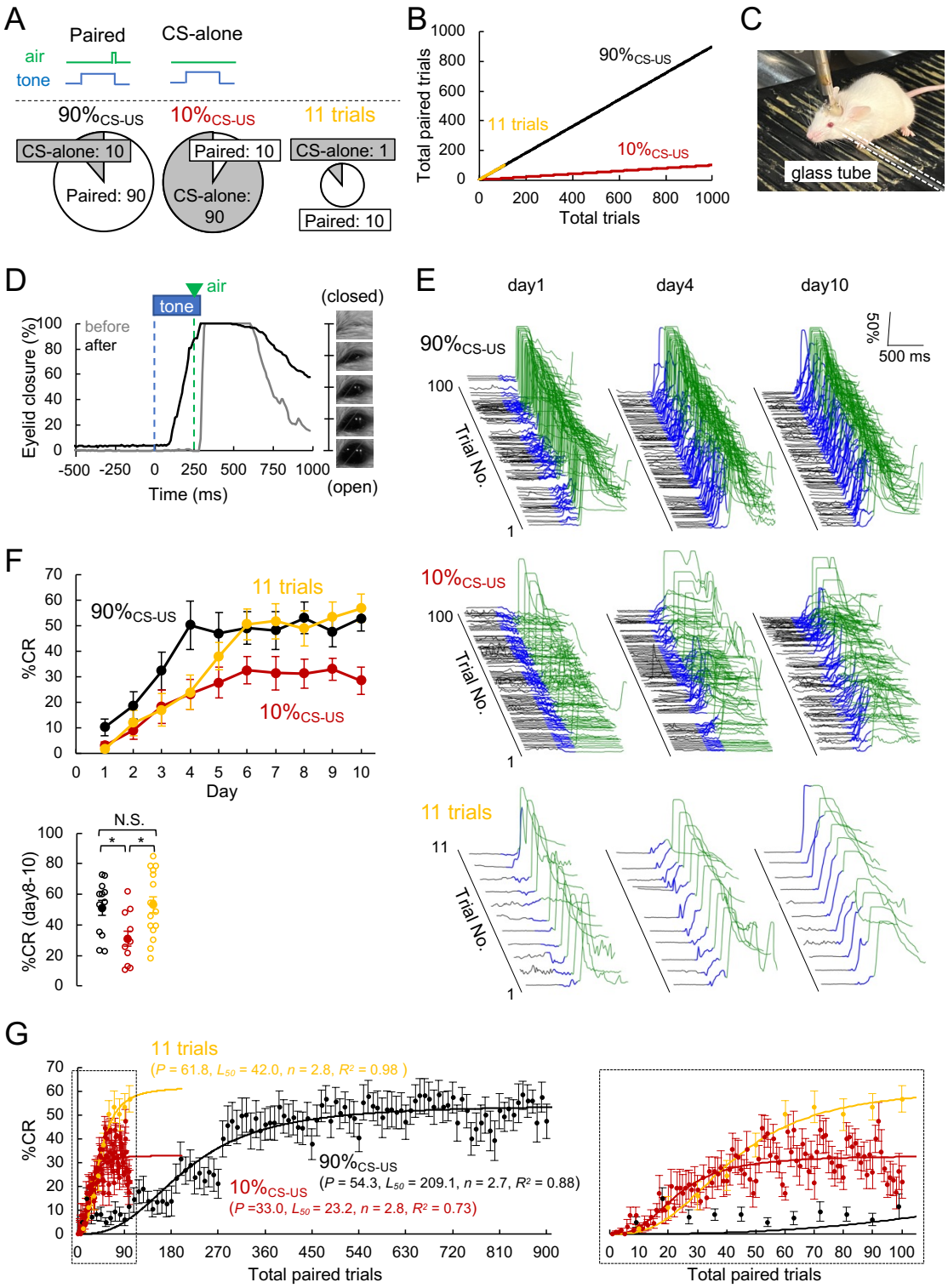

Figure 1. Eyeblink conditioning with distinct patterns of sensory pairings

A, Schematics of CS (tone)-US (air puff) and CS-alone trials (top) and three training protocols, 90%$_{CS\text{-}US}$, 10%$_{CS\text{-}US}$, and 11 trials (bottom). **B**, Total paired trials number plotted against total trials during 10 days training (90%$_{CS\text{-}US}$: black; 10%$_{CS\text{-}US}$: red; 11 trials: orange). **C**, An image of a head-

fixed mouse on a treadmill during experiments. A glass tube put in front of the eyelid is outlined by white dashed line. **D**, Representative images and time courses of eyelid closure before/after the CR acquisition trainings. **E**, Representative eyelid traces (100 trials from a single mouse) for the $90\%_{CS\text{-}US}$ (top), $10\%_{CS\text{-}US}$ (middle) and 11 trials (bottom) on day1 (left), 4 (middle) and 10 (right). **F**, Daily changes of %CR established by 10 days of training (top) and averaged %CR during day8-10 (bottom; $90\%_{CS\text{-}US}$: 51.1 ± 5.0%, n = 12; $10\%_{CS\text{-}US}$: 30.9 ± 4.9%, n = 11; 11 trials: 53.0 ± 5.1%, n = 16, Tukey-Kramer test), averaged from data showing CR establishment (mice showing CR ≥ 10% of trials in at least 3 days). Individual, open circle; average, filled circle. **G**, %CR averaged from each 10 or 11 trials plotted against the number of total paired trials ($90\%_{CS\text{-}US}$: n = 18; $10\%_{CS\text{-}US}$: n = 13; 11 trials: n = 17). Fitted curves are also shown.

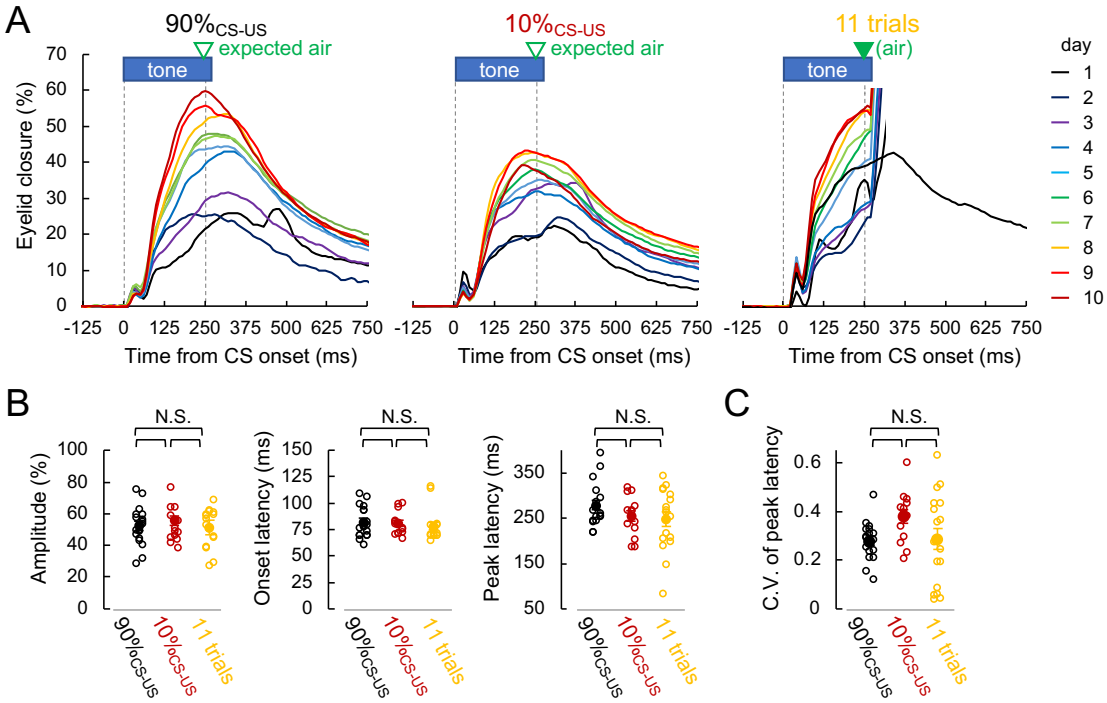

Figure 2. Learned eyelid responses

A, Time courses of averaged eyelid closure for each day (90%$_{CS-US}$: n = 18 obtained from CS-alone trials; 10%$_{CS-US}$: n = 13 from CS-alone trials; 11 trials: n = 17 from paired & CS-alone trials). For 11 trials, a black trace obtained from averaging eyelid closures in CS-alone trials during 10 days is also shown. For averaging, eyelid traces showing responses were used. Baseline level before CS was aligned to 0% for each. **B**, **C**, Averaged amplitude (left), onset latency (middle) and peak latency (right) of eyeblink responses (B), and coefficient of variance for peak latency (C) for all sessions. Individual, open circle; average, filled circle. Tukey-Kramer test.

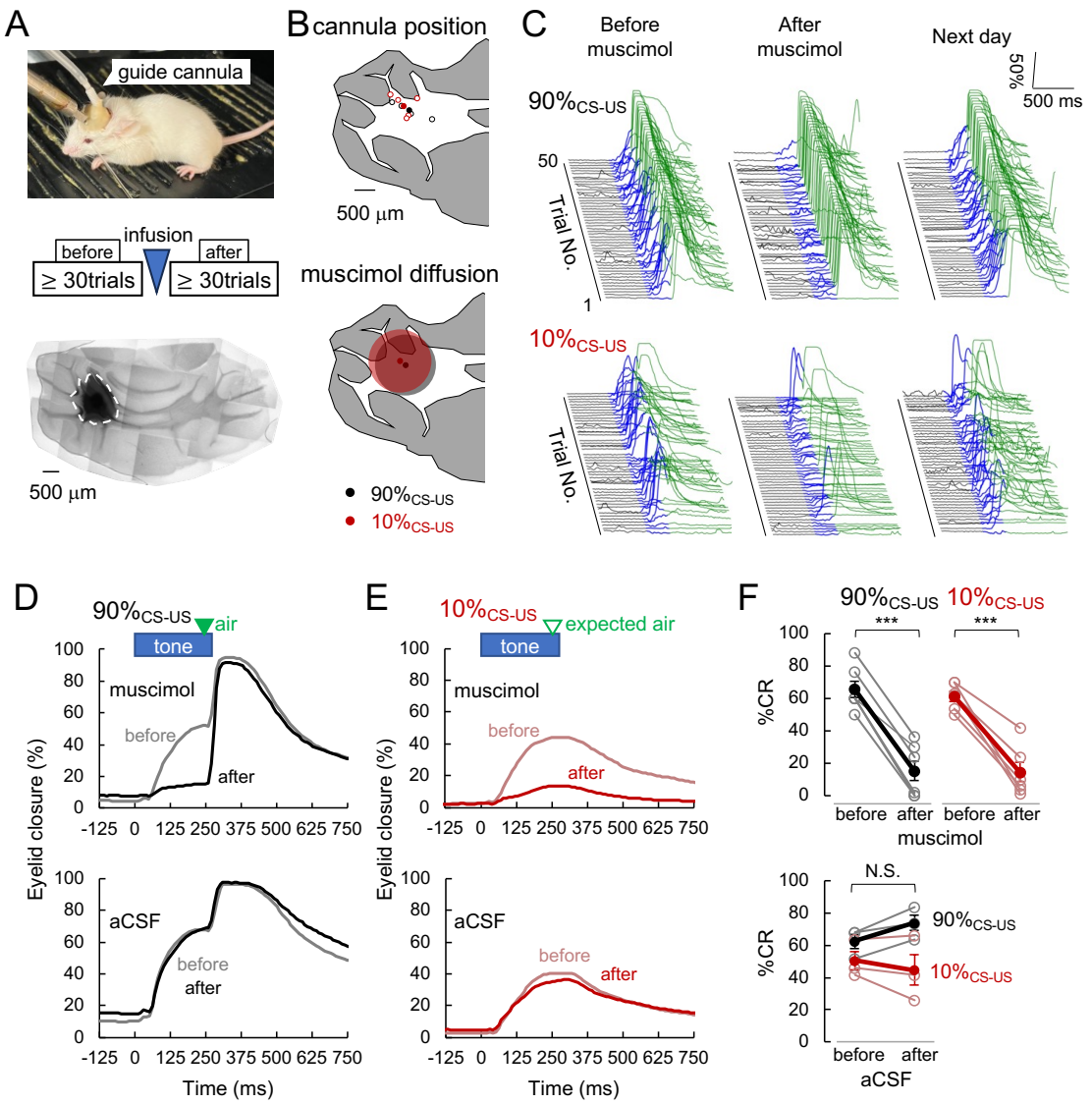

Figure 3. Critical role of DCN activity for occurrence of established CRs

A, **B** Schematics of pharmacological inactivation by muscimol (A) and the target location (B) in the cerebellum. Muscimol or aCSF was infused through an implanted guide cannula (top). Mice had at least 30 trials before/after muscimol or aCSF infusion (middle). Area for muscimol diffusion identified from fluorescence on a horizontal cryosection is shown in A, bottom (white dashed line), and summarized in B (90%$_{CS\text{-}US}$, n = 5; 10%$_{CS\text{-}US}$, n = 5). Cannula tip positions are also shown in B, top (90%$_{CS\text{-}US}$, n = 4; 10%$_{CS\text{-}US}$, n = 5). The distance of each average cannula position was less than 150 μm and the diameter of muscimol diffused area was 1.45 ± 0.19 mm for the 90%$_{CS\text{-}US}$ group; 1.50 ± 0.12 mm for the 10%$_{CS\text{-}US}$ group. **C**, Representative eyelid traces from a single mouse in the 90%$_{CS\text{-}US}$ (top) and 10%$_{CS\text{-}US}$ (bottom) before, after and the next day of muscimol infusion. **D**, **E**, Averaged eyelid traces before/after muscimol (top) or aCSF (bottom)

infusion in the $90\%_{CS\text{-}US}$ (D, n = 6) and $10\%_{CS\text{-}US}$ (E, n = 6). Eyelid traces from all acceptable trials were used for averaging. **F**, %CR changes before/after infusion of muscimol (top) and aCSF (bottom). Paired t-test.

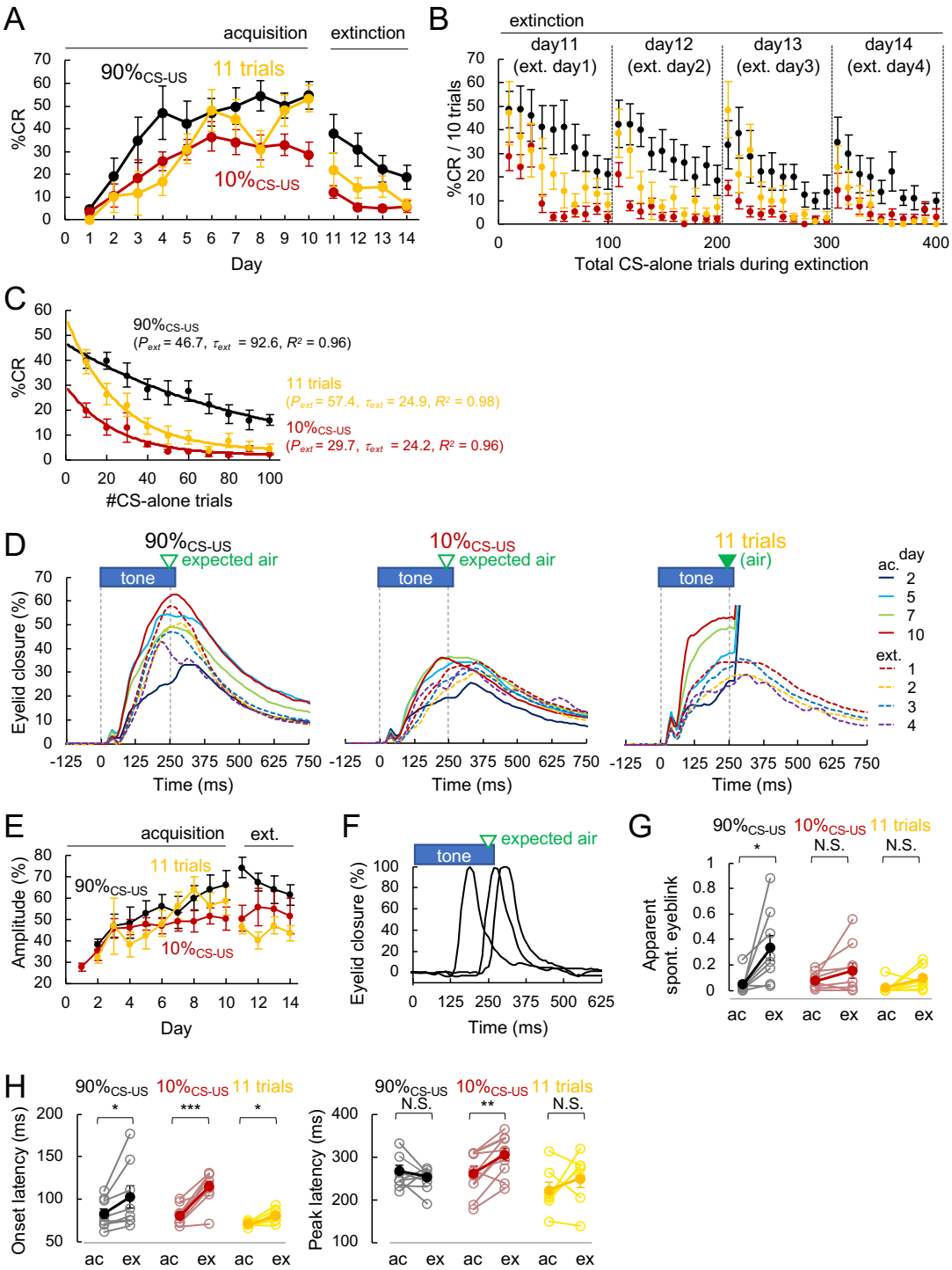

Figure 4. Dynamics of CR probability and eyelid closures during memory extinction

A, **B**, Daily %CR changes during 14 days of training consisting of acquisition and extinction sessions (A) and %CR averaged from each 10 trial during 4 days extinction (B). 90%$_{CS\text{-}US}$, n = 8; 10%$_{CS\text{-}US}$, n = 9; 11 trials, n = 7. **C**, Averaged %CR changes of every 10 trial from the beginning of extinction day1-4 were plotted against the number of CS-alone trials with fitting by single exponential curves. **D**, **E,** Averaged eyelid closure traces on acquisition day2, 5, 7, 10 and

extinction day1-4 (D), indicated by different colors, and daily changes of averaged amplitude of eyelid closure (E) during acquisition and extinction paradigms. For the $90\%_{CS\text{-}US}$ and $10\%_{CS\text{-}US}$ from CS-alone trials; for the 11 trials from paired & CS-alone trials. The baseline level before CS was aligned to 0% for each. **F**, Example eyelid traces of responses like spontaneous eyeblink (> 90% of eyelid closure with half-width < 150 ms). **G**, Ratio of spontaneous eyeblink-like ones to all responses. Acquisition, ac; extinction, ex. Paired t-test. **H**, Onset latency (left) and peak latency (right) during acquisition and extinction training. Paired t-test.

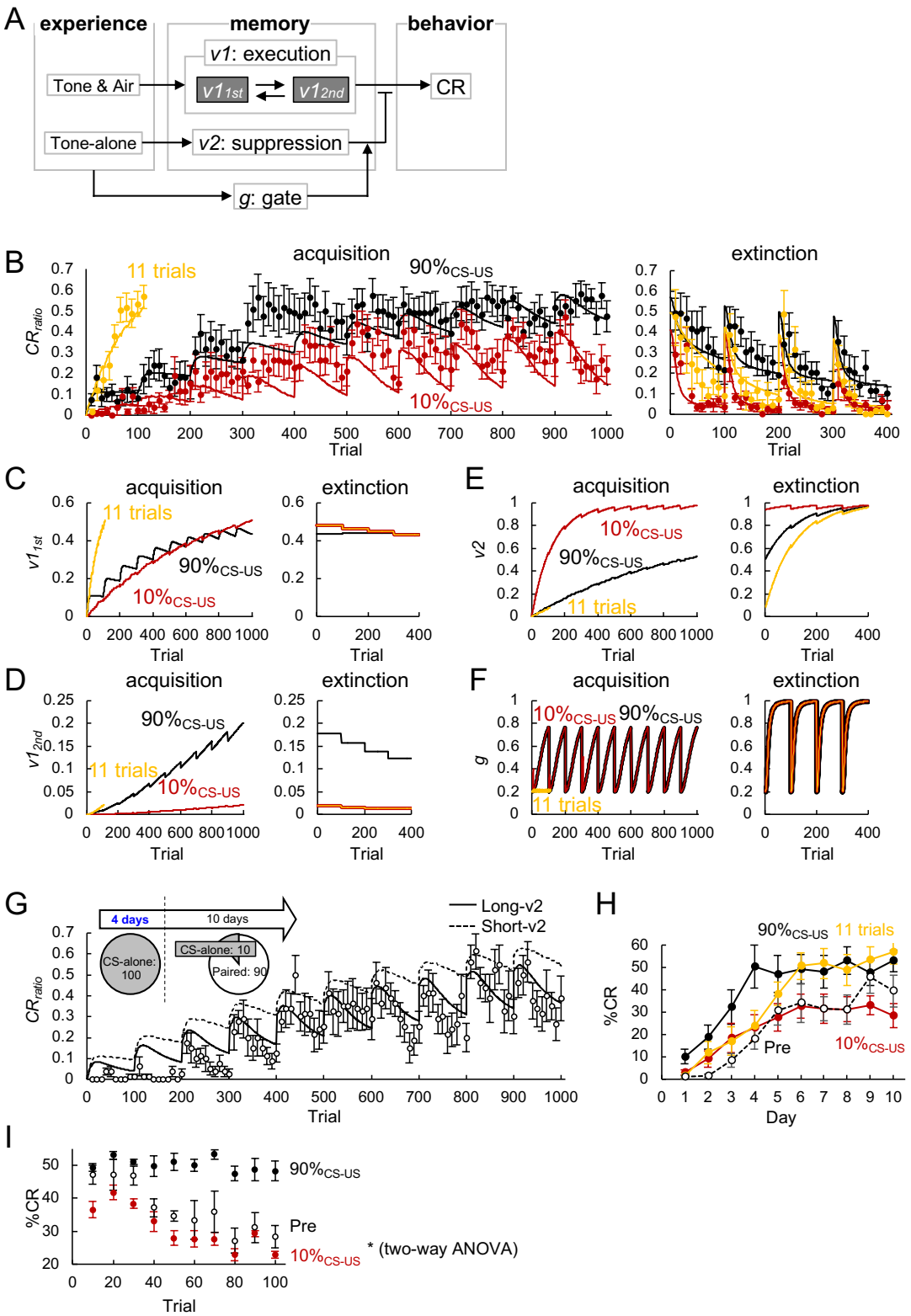
A
experience
memory
behavior
v1: execution
Tone & Air
v11st
v12nd
CR
Tone-alone
v2: suppression
g: gate
B
acquisition
extinction
11 trials
90%CS-US
10%CS-US
CRratio
Trial
C
acquisition
extinction
v11st
11 trials
90%CS-US
10%CS-US
Trial
D
acquisition
extinction
v12nd
90%CS-US
11 trials
10%CS-US
Trial
E
acquisition
extinction
v2
10%CS-US
90%CS-US
11 trials
Trial
F
acquisition
extinction
g
10%CS-US
90%CS-US
11 trials
Trial
G
4 days
10 days
CS-alone: 100
CS-alone: 10
Paired: 90
Long-v2
Short-v2
CRratio
Trial
H
90%CS-US
11 trials
Pre
10%CS-US
%CR
Day
I
90%CS-US
Pre
10%CS-US
* (two-way ANOVA)
%CR
Trial

Figure 5. Modeling of formation of counteracting memories for CRs

A, Model scheme for input-dependent memory formation for CR execution and its suppression. **B**, Simulated CR_{ratio} plotted as lines against total trial number during acquisition (left) and extinction (right) paradigms together with CR ratio averaged from each 10 or 11 trial recorded in experiments (dots, $90\%_{CS\text{-}US}$, n = 12; $10\%_{CS\text{-}US}$, n = 11; 11 trials, n = 16). **C-F**, Two components for memory to execute CR, $v1_{1st}$ (C) and $v1_{2nd}$ (D), the counteracting memory to suppress CR, $v2$ (E), and its gate g (F) are plotted against total trial number during acquisition (left) and extinction (right) in simulation. **G**, Prediction of CR_{ratio} changes during acquisition by $90\%_{CS\text{-}US}$ following the 4 days CS-alone exposure (Pre) in models assuming long (solid line)- or short (dotted line)-lasting $v2$. Actual experimental results of CR ratio averaged from each 10 trial (n = 8, dot) are also shown to evaluate the model prediction. Training protocols for the Pre group (top left). **H**, Daily %CR changes during 10 days of acquisition training. Pre, n = 8. Data for the $90\%_{CS\text{-}US}$ (n = 12), $10\%_{CS\text{-}US}$ (n = 11), and 11 trials (n = 16) shown in Fig. 1F are also superimposed for comparison. **I**, Averaged %CR changes of every 10 trial from the beginning of day7-10 were plotted against trials number. A two-way repeated measures of analysis of variance (ANOVA) found an interaction of experience patterns and trials number to %CR.