Multiple sites of extinction for a single learned response

Brian E. Kalmbach1 and Michael D. Mauk1,2

1Center for Learning and Memory, 2Section of Neurobiology, University of Texas, Austin, Texas

Submitted 25 April 2011; accepted in final form 19 September 2011

Kalmbach BE, Mauk MD. Multiple sites of extinction for a single learned response. J Neurophysiol 107: 226–238, 2012. First published September 21, 2011; doi:10.1152/jn.00381.2011.—Most learned responses can be diminished by extinction, a process that can be engaged when a conditioned stimulus (CS) is presented but not reinforced. We present evidence that plasticity in at least two brain regions can mediate extinction of responses produced by trace eyelid conditioning, where the CS and the reinforcing stimulus are separated by a stimulus-free interval. We observed individual differences in the effects of blocking extinction mechanisms in the cerebellum, the structure that, along with several forebrain structures, mediates acquisition of trace eyelid responses; in some rabbits extinction was prevented, whereas in others it was largely unaffected. We also show that cerebellar mechanisms can mediate extinction when noncerebellar mechanisms are bypassed. Together, these observations indicate that trace eyelid responses can be extinguished via processes operating at more than one site, one in the cerebellum and one upstream in forebrain. The relative contributions of these sites may vary from animal to animal and situation to situation.

Pavlovian; nictitating; eyelid; plasticity; learning; delay; trace; cerebellum; prefrontal cortex

Behavioral extinction is an important aspect of learning. Learned responses that were once appropriate may become inappropriate in other contexts or as circumstances change, and the mechanisms of extinction are relevant to a number of psychiatric disorders (Peters et al. 2009; Phelps and LeDoux 2005). As with the mechanisms of acquisition, identifying the sites of plasticity is an essential initial step toward understanding the mechanisms of extinction. Eyelid conditioning is an example of the many forms of learning that require more than one site of plasticity (Medina et al. 2002b). For example, delay eyelid conditioning involves plasticity in at least two sites within the cerebellum (Ohyama et al. 2006; Perrett et al. 1993). Such instances raise interesting and potentially complex possibilities for the site(s) of plasticity underlying extinction. If plasticity at more than one site is necessary for the expression of learned responses, reversing or counteracting plasticity at any site may be sufficient to mediate extinction. Moreover, multiple sites of plasticity, each potentially sufficient to support extinction, create the potential for animal-to-animal and situation-to-situation variability in the site or sites that mediate extinction. We report evidence that at least two different sites of plasticity can contribute to the extinction of trace eyelid responses in rabbit.

Eyelid conditioning provides many advantages for addressing the mechanisms of extinction. Delay eyelid conditioning, where pairing a conditioned stimulus (CS; often a tone) with a reinforcing unconditioned stimulus (US) promotes learned eyelid closure to the CS, is mediated by plasticity in both the cerebellar cortex and deep cerebellar nuclei (DCN) (Medina et al. 2001; Ohyama and Mauk 2001; Ohyama et al. 2006; Perrett et al. 1993). This plasticity is engaged by pairing tone-driven mossy fiber inputs (Halverson and Freeman 2010; Steinmetz et al. 1987) with US-driven climbing fiber inputs to the cerebellum (Mauk et al. 1986; Rosen et al. 1989; Yeo et al. 1986). Evidence indicates that delay responses extinguish via plasticity in cerebellar cortex, leaving plasticity in the DCN somewhat intact (Fig. 1) (Jirenhed et al. 2007; Medina et al. 2001, 2002a; Perrett and Mauk 1995).

Although delay eyelid conditioning is largely independent of forebrain structures (Mauk and Thompson 1987; Oakley and Russell 1972), trace eyelid conditioning, in which a stimulus-free trace interval separates CS offset and US onset, requires several forebrain structures for trace intervals longer than ~300–400 ms and engages cerebellar learning via a mossy fiber input driven by medial prefrontal cortex (mPFC) in addition to the tone-driven mossy fiber input necessary for delay conditioning (Fig. 1) (Fontan-Lozano et al. 2005; Kalmbach et al. 2009, 2010a, 2010b; Mauk and Thompson 1987; Moyer et al. 1990; Powell and Churchwell 2002; Ryoo et al. 2001; Takehara et al. 2003; Weible et al. 2000, 2007). This highlights the possibility that both cerebellar and forebrain sites of plasticity could support extinction in trace eyelid conditioning.

We report evidence that the extinction of trace eyelid responses can indeed be mediated by more than one site, one in the cerebellum and another presumably in upstream forebrain structures. We also found evidence that the site of extinction can vary from animal to animal. These results demonstrate that extinction of a single learned response can be mediated by plasticity at more than one site. The contributions from these sites apparently can vary depending on the situation or learning regimen, and this may even explain the bimodal distribution of extinction rates observed during extinction.

Materials and Methods

Subjects and surgery. Male New Zealand albino rabbits (Oryctolagus cuniculus; Myrtle’s Rabbitry, Thompsons Station, TN) initially weighing 2.5–3 kg served as subjects. Rabbits were individually housed, maintained on a fixed daily diet, and given free access to water. All surgical and experimental procedures were approved by The University of Texas at Austin Institutional Animal Care and Use Committee and were in accordance with the National Institutes of Health guidelines.

Rabbits were prepared for microstimulation and/or infusion experiments using sterile surgical procedures. An initial anesthetic cocktail of ketamine (40 mg/kg) and acepromazine (5 mg/kg) was delivered via subdermal injection. Rabbits were then placed in a stereotaxic restrainer, and general anesthesia was maintained with isoflurane (2% mixed in oxygen) for the remainder of the surgery. An incision of ~4 cm was made along midline, and the skin was retracted to

Address for reprint requests and other correspondence: B. E. Kalmbach, Center for Learning and Memory, Section of Neurobiology, Univ. of Texas, 1 Univ. Station, C7000, Austin, TX 78712 (e-mail: brian@mail.clm.utexas.edu).
reveal the lambda and bregma landmarks of the skull. A 3-mm-diameter craniotomy was then drilled to accommodate the cannula and/or stimulation electrodes, and four other smaller craniotomies were drilled to accommodate anchor screws. Lambda was then positioned 1.5 mm ventral to bregma. In six subjects, a stainless steel guide cannula (Plastics One, Roanoke, VA) was targeted for the right dorsal accessory olive (1.0 mm anterior, 0.7 mm lateral, and 22 mm ventral from lambda). In eight subjects, a guide cannula was targeted for the left anterior interpositus nucleus of the DCN (1 mm anterior, 5 mm lateral, and 13.3 mm ventral from lambda). In five subjects, two closely spaced tungsten stimulating electrodes laterally separated by 1.0 mm (A-M Systems, Carlsborg, WA; ~100–200 kΩ) were placed in the left middle cerebellar peduncle (3.0 mm anterior, 5.0 mm lateral, and 16.0 mm ventral), and a guide cannula was placed in the interpositus nucleus. In these subjects, a stainless steel screw placed on the surface of the brain served as ground. In all subjects, including 21 subjects used only for behavioral experiments, a bolt was placed between the skull screws and dental acrylic was used to secure the implants in place. Loose skin was sutured, and a dummy cannula was placed in the guide cannula. Finally, stainless steel electrodes were placed at the margins of the left eye (one rostral and the other caudal) to deliver the electrical stimulation US. Rabbits were given postoperative analgesics and antibiotics for 2 days and were allowed to recover for a week before experiments began.

**Conditioning.** Subjects were trained in custom-designed, well-ventilated, and sound-attenuating chambers measuring ~90 × 60 × 60 cm (length, width, height). To generate tones for the CS and electrical pulses for the US, each chamber was equipped with a speaker connected to an audio source module (model V85-05; Coulborn Instruments, Allentown, PA) and with isolated pulse stimulators (model 2100; A-M Systems, Carlsborg, WA) connected via leads to the electrodes implanted around the eye. Other stimulus timers were used to time the delivery of constant-current pulses by stimulus isolators (model 2300; A-M Systems) through leads connected to electrodes implanted in the middle cerebellar peduncle. The position of the left external eyelid was measured by using an infrared emitter/detector to detect changes in the amount of reflected light as the eyelids open and close. At the start of each daily conditioning/test session, the eyelid detector was attached to the bolt on the head stage of each rabbit and calibrated by delivering the US to elicit maximum eyelid closure. The amplitude of the signal was adjusted to match an assumed 6-mm maximum eyelid closure.

Stimulus presentation and data acquisition were controlled by custom-designed software that was run on a computer adjacent to the conditioning chambers. Subjects received trace conditioning, dual delay/trace conditioning, or conditioning with mossy fiber stimulation as the CS. All daily conditioning sessions consisted of 12 blocks of 9 trials where the first trial of each block was the CS presented alone. Trials were separated by a mean interval of 30 s (±10 s).

Twelve rabbits were trained for 10–12 days in trace conditioning with a 1-kHz sinusoidal tone (ramped at onset and offset with a time constant of 5 ms to avoid clicks). The US was a 50-ms train of constant-current pulses (50 Hz, 1-ms pulse width, 2–3 mA) delivered through electrodes implanted around the left eye. For paired trace conditioning trials, the CS was presented for 500 ms and was followed 500 ms later by the US.

Fourteen rabbits were trained for 10–15 days using a dual delay/trace conditioning paradigm in which delay and trace conditioning trials alternated. The tone (1 or 9.5 kHz) used to signal each trial type was counterbalanced across subjects. Trace conditioning trials were the same as described above, whereas for paired delay conditioning trials the 50-ms US was presented 500 ms into a 550-ms CS.

Five rabbits were trained with electrical stimulation of mossy fibers as the CS. Two stimulus trains with durations that approximate tone-driven (Aitkin and Boyd 1978; Boyd and Aitkin 1976; Campolattaro et al. 2011; Freeman and Muckler 2003) and putative mPFC-driven mossy fiber inputs during trace conditioning (Takehara-Nishiuchi and McNaughton 2008) were used, with a 500-ms CS and 500-ms trace interval delivered through two separate electrodes implanted in the middle cerebellar peduncle (where mossy fibers enter the cerebellum). This pattern of input has previously been shown to be necessary and sufficient to produce responses that quantitatively match trace conditioned responses to a tone (Kalmbach et al. 2010a). Stimulus trains consisted of 100-μs, 100-μA constant-current pulses delivered at 50 Hz. The tone-mimicking input lasted for 500 ms, and the mPFC-mimicking input lasted for 1,000 ms. Training sessions were otherwise identical to those described above. Rather than train subjects for a set number of days, subjects were trained one session past the session in which they first reached a 60% response rate. We stimulated mossy fibers rather than their cell bodies in the pontine nuclei to minimize activation of fibers of passage or antidromic activation of pontine-projecting forebrain regions. All electrodes were placed within the middle cerebellar peduncle (~2–3 mm anterior to the anterior interpositus nucleus as confirmed by histological analysis. Furthermore, stimulation through each electrode did not elicit any noticeable movements. Extinction sessions were identical to training sessions except that all trials were CS-alone presentations.

**Test sessions and infusions.** After these initial training sessions, a subset of subjects were given extinction sessions where infusions were made in either the DCN (8 dual delay/trace-trained and 5 stimulation-trained subjects) or the dorsal accessory olive (6 dual delay/trace-trained subjects). Drugs were dissolved in artificial cerebrospinal fluid (ACSF) consisting of (in mM) 124 NaCl, 3.0 KCl, 26.0 NaHCO₃, 1.3 NaH₂PO₄·H₂O, 2.0 MgCl₂, 10.0 dextrose, 10.0 HEPES (pH 7.35), and 2.0 CaCl₂. Infusions were made through a 33-gauge infusion cannula that extended 1.2 mm beyond the surgically implanted guide cannula. The infusion cannula was coupled to a 50-μl Hamilton syringe that was mounted on an automated injector system.
In subjects with the cannula in the DCN (8 tone-trained and 5 stimulation-trained subjects), there were two infusion sessions (Fig. 2). Twenty minutes before the start of the extinction session, ACSF or the GABA agonist muscimol (1 mM; Tocris) was infused into the DCN at a rate of 0.1 μl/min (total volumes 7–8 μl). The infusion then continued for the remainder of the session at the same rate. An additional day of extinction was given after each infusion session, and subjects were retrained between infusion sessions (each subject received both types of infusions on separate days). For stimulation-trained subjects, the infusion session administered first was counterbalanced across subjects.

Subjects with the cannula in the inferior olive were administered two types of infusion sessions. Twenty minutes before the start of the extinction session, ACSF or the GABA antagonist gabazine (20 μM; Tocris) was infused into the inferior olive at a rate of 0.1 μl/min (total volumes 7–8 μl). The infusion then continued for the remainder of the session at the same rate. All subjects were given each type of infusion, and the type of infusion given first was counterbalanced across subjects. If gabazine infusions did not affect delay responding, subjects were retrained and the infusion session was repeated with a 0.5-mm longer infusion cannula. Two subjects were given an additional infusion session where the AMPA receptor antagonist 2,3-dihydroxy-6-nitro-7-sulfamoyl-benzo(f)quinoxaline (NBQX; 150 μM) was infused into the olive at a rate of 0.1 μl/min. In one subject, the infusion began 20 min before the start of the session, and in the other subject, the infusion began 40 min before the start of the session. Each infusion continued for the remainder of the session at a rate of 0.1 μl/min.

**Data analysis.** Eyelid responses during each trial were digitized (1 kHz, 12-bit resolution) and stored on disk for subsequent off-line analyses using custom software. For each response, 2,500 points were recorded, the 200 ms before CS onset and the 2,300 ms that followed. Trials in which eyelid movements >0.3 mm were made in the 200 ms before CS onset were automatically excluded from analysis by the software (fewer than 5%). A conditioned response during paired conditioning sessions was defined as an eyelid movement of at least 0.3 mm within the interval between CS onset and US onset (inter-stimulus interval; ISI). For extinction training, this interval was extended 200 ms beyond the training ISI. Latency to onset was determined using an algorithm designed to detect the initial deflection of each response away from the baseline. The criterion for extinction was defined as eight non-conditioned responses in nine consecutive trials. Thus trials to criterion is the total number of trials performed to reach this criterion.

Repeated-measures analysis of variance (ANOVA) followed by F-tests for simple effects and Tukey’s post hoc comparisons were used to test for within-subject differences. The Kolmogorov-Smirnov test was used to test for normality. All tests were two tailed with a significance level of 0.05.

**Histology.** After the completion of experiments, marking lesions were made by passing DC anodal current (200 μA) for 20 s through the electrodes or a stainless steel wire cut the length of the guide cannula. Animals were killed with an overdose of pentobarbital sodium and perfused transcardially with 1 liter of 0.9% saline followed by 1 liter of 10% formalin. Brains were extracted and stored in 10% formalin for at least a week. They were then embedded in an albumin gelatin mixture and sectioned using a freezing microtome (80-μm sections). Tissue was mounted and stained with cresyl violet.

**RESULTS**

**Disrupting cerebellar mechanisms of extinction prevents extinction of delay but not trace responses.** Previous work has shown that inhibition of the climbing fiber input to the cerebellum via a projection from DCN to the inferior olivary nuclei (Bengtsson et al. 2004; Hesslow and Ivarsson 1996) is the necessary and sufficient signal for extinction of delay eyelid responses (Medina et al. 2002a). Different methods of blocking this inhibitory signal all prevent the extinction of delay eyelid responses (Hardiman et al. 1996; Medina et al. 2002a; Ramnani and Yeo 1996). We began by testing whether blocking this inhibitory signal, and thus blocking the cerebellar mechanisms of extinction, prevents extinction of trace eyelid responses. We employed both methods known to prevent extinction of delay eyelid responses: infusing a GABA_A receptor agonist into the DCN to block activity of DCN neurons (Fig. 1a) (Hardiman et al. 1996; Ramnani and Yeo 1996) and infusing a GABA_A receptor antagonist into the inferior olive (Fig. 1b) (Medina et al., 2002a). We implemented these experiments in rabbits trained with both delay and trace conditioning to permit within-animal comparison of the effects on extinction of delay and trace responses.

**Preventing cerebellum-mediated extinction by silencing neurons in DCN.** Silencing DCN neurons with infusions of the GABA_A receptor agonist muscimol not only prevents DCN-mediated inhibition of the climbing fibers during the CS but also prevents the expression of conditioned responses (Chapman et al. 1990; Garcia and Mauk 1998; Hardiman et al. 1996; Kalmbach et al. 2009; Ramnani and Yeo 1996). Because of this, the ability of a muscimol infusion to block extinction must be inferred from levels of responding during the postmuscimol extinction session the next day (Ramnani and Yeo 1996). In essence, if responding during this second postmuscimol extinction...
tion session is similar to that during a typical first session of extinction, then the previous day’s infusion of muscimol blocked extinction. In contrast, lower levels of responding more typical of a second session of extinction training indicate the infusion did not block extinction (Fig. 2A).

To make these comparisons for both delay and trace conditioning, we trained eight rabbits using a dual delay/trace conditioning procedure until an asymptotic level of responding was reached. In this procedure, one tone CS was used with delay conditioning and a second tone CS was used with trace conditioning, and for all analyses the delay and trace trials were analyzed separately for each animal. Each rabbit was then given a standardized set of five sessions (Fig. 2A): 1) extinction during infusion of muscimol to block cerebellar mechanisms of extinction, 2) a postmuscimol extinction session, 3) retraining back to asymptotic performance, 4) re-extinction during a control infusion of the ACSF vehicle (control day 1), and 5) a second (control day 2) extinction session with no infusion. The latter 2 days (control 1 and control 2) serve as the control levels of responding on the first and second days of normal extinction. Again, the key comparisons are the rates of extinction on the postmuscimol session vs. the control 1 session and vs. the control 2 session. These within-animal comparisons are possible, because previous studies have shown that repeated rounds of extinction and reacquisition do not appreciably change the rate of trace response extinction (Kehoe 2006).

It is essential to the interpretation of this experiment that the infusions successfully abolished the expression of conditioned responses during the muscimol infusion session. To test the effectiveness of these infusions, we compared the rate of responding during the muscimol infusion session to the rate of responding during the later ACSF infusion session. Figure 3 shows that the infusion of muscimol into the DCN abolished both delay [32.7 ± 5.0% responding during ACSF infusion, 5.6 ± 2.6% responding during muscimol infusion; t(7) = 5.26, P = 0.001] and trace eyelid responses [12.86 ± 2.2% responding during ACSF infusion, 3.5 ± 1.1% responding during muscimol infusion; t(7) = 4.46, P = 0.003]. Sample cannula placements in the DCN are shown in Fig. 3B.

Comparison of performance during the subsequent postmuscimol extinction session with the control 1 and control 2 sessions revealed that the muscimol infusions prevented the extinction of delay but not trace eyelid responses. For delay conditioning, the level of responding during the entire postmuscimol extinction session was indistinguishable from that during the control 1 session (P = 0.63; paired t-test) and was significantly greater than that during the control 2 session (t(7) = 4.58, P = 0.003; paired t-test). We also examined the rate of extinction for these groups with a two-way repeated-measures ANOVA. Figure 4A shows that delay responses during postmuscimol and control 1 sessions extinguished at similar rates (compare dark gray and light gray curves in Fig. 4A; P = 0.32; block × session interaction effect), whereas responses during the control 2 session extinguished more quickly [black curve; F(1, 77) = 2.27, P = 0.02 for postmuscimol and F(1, 77) = 6.15, P < 0.001 for ACSF (control 1); block × session interaction effects]. Thus, in terms of overall responding for the entire session and in terms of rate of extinction, the postmuscimol delay responding was like a first (control 1) day of extinction and not like a second (control 2) day of extinction. These results indicate that, for the delay responses, the process of extinction was blocked during the muscimol infusion. Conversely, for the trace responses, the muscimol infusion, control 1, and control 2 sessions did not differ in terms of the overall level of responding during the entire session, but the postmuscimol and control 1 sessions showed significantly different rates of extinction (Fig. 4, A and C). Responses during the postmuscimol and control 2 sessions extinguished at the same rate (compare dark gray and black curves in Fig. 4A; P = 0.94 for block × session interaction effect), whereas extinction during the postmuscimol session occurred more quickly than for the control 1 session [light gray curve; F(1, 77) = 5.02, P < 0.001 for postmuscimol block × session interaction effect]. Thus the postmuscimol trace responding was different from a first (control 1) day of extinction and indistinguishable from a normal second (control 2) day of extinction. These results indicate that the muscimol infusion had no measurable effect on the extinction of the trace responses.

This differential effect of muscimol infusions on delay vs. trace extinction is clearly apparent in the session difference scores obtained by subtracting (for each animal) the response likelihood of each training block during the postmuscimol extinction session from the same block of either the control 1 session or the control 2 session. A score of zero indicates that performance during the two sessions under comparison is equal for a given training block. For delay conditioning, response likelihood during postmuscimol sessions was approximately equal to control 1 sessions (difference scores ≈ 0), whereas response likelihood during postmuscimol extinction was greater than control 2 extinction for the first few training blocks (difference scores < 0; Fig. 4B, left). For trace conditioning, the inverse pattern of results is apparent; response likelihood during postmuscimol and control 2 extinction was
approximately equal across training blocks (difference scores \( \approx 0 \)), whereas response likelihood during postmuscimol extinction was less than the control 1 session for the first few training blocks (difference scores > 0; Fig. 4B, right). Together, these data suggest that inactivation of DCN, and thus blocking cerebellum-mediated inhibition of climbing fibers, prevents the extinction of delay eyelid responses, which replicates the findings of previous studies (Hardiman et al. 1996; Ramnani and Yeo 1996), but does not prevent the extinction of trace eyelid responses.

**Preventing cerebellum-mediated extinction by blocking inhibition of climbing fibers.** Blocking the inhibition of climbing fibers directly via infusion of the GABA\(_A\) receptor antagonist gabazine into the inferior olive (Medina et al. 2002a) represents a second and more direct way to interrupt extinction mechanisms in the cerebellum. This procedure has the advantage of not blocking the expression of conditioned responses and thus allows any effects on extinction to be observed directly during the infusion session. Six rabbits were again trained using the dual delay/trace procedure to asymptotic levels of responding to both trial types. Each rabbit was then given two test sessions, separated by a day of retraining, to test the effects of gabazine or ACSF on extinction. During these test sessions, gabazine (20 \( \mu \)M) or ACSF was infused into the inferior olive beginning 20 min before the start of extinction training. As before, delay and trace trials were analyzed separately with effects on delay responses serving as an indication that extinction in the cerebellum was blocked by the infusion. Similar to the effects of muscimol infusions in the DCN, gabazine infusions in the inferior olive prevented the extinction of delay but not trace eyelid responses. For delay conditioning, the average response likelihood during the gabazine infusion was greater than during the ACSF infusion session [Fig. 5, A, C, left, and D, left; \( F_{(1,5)} = 64.20, P < 0.001 \); sessions effect].
whereas it did during the ACSF infusion. The eyelid position sweeps in Fig. 6 are from four of these six rabbits and demonstrate the range of effects of the gabazine infusion. In addition, the average response likelihood for each rabbit during both the gabazine and ACSF infusion sessions is shown in Fig. 6. Differences in cannula placements do not appear to account for these differences, because cannula were similarly placed in the anterior portion of the inferior olive (Fig. 6).

This rabbit-to-rabbit variability in the ability of the gabazine infusion to prevent trace extinction, in the face of the ability of the same infusions to consistently prevent delay extinction, suggests that extinction of trace eyelid responses can be mediated by plasticity at more than one site. One site requires inhibition of climbing fiber inputs, similar to the mechanism that mediates extinction of delay eyelid responses, and the other mechanism appears to involve plasticity outside of the cerebellum. On the basis of previous work demonstrating that trace eyelid conditioning requires an input to the cerebellum driven by the prefrontal cortex (Kalmbach et al. 2009) (Fig. 1), we hypothesize that extinction can be mediated by plasticity that removes or shortens the duration of this input. Conversely, when this mechanism fails or occurs relatively slowly, trace extinction can occur by the cerebellar mechanism that involves inhibition of climbing fiber inputs. As described below, we tested two key corollaries of this hypothesis: that cerebellar mechanisms are sufficient to mediate extinction of trace eyelid responses (when the noncerebellar mechanism is bypassed) and that there is a bimodal distribution in the rate of trace response extinction, consistent with the potential contributions of two different extinction mechanisms.

Cerebellar mechanisms can support extinction of trace eyelid responses. Testing the hypothesis that cerebellar mechanisms are sufficient for extinction of trace eyelid responses requires a way to test extinction when the noncerebellar mechanism is bypassed. We employed two methods to accomplish this. The first is based on evidence that blocking the ability of the US to activate climbing fibers (Fig. 1C) induces cerebellum-mediated extinction, even during paired presentations of the CS and US (McCormick et al. 1985; Medina et al. 2002a). Thus, by pharmacologically blocking the US input to the cerebellum during paired training, we can induce cerebellum-mediated extinction while presumably any noncerebellar mechanism is still exposed to the CS and US in each trial. Two subjects that were previously infused with gabazine in the inferior olive were retested using infusions of the AMPA receptor antagonist NBQX during paired presentations of the CS and US. This infusion blocks the US input to the cerebellum, which has been shown previously to induce extinction of delay responses even when the CS is paired with the US (Medina et al. 2002a) but presumably does not block US input to other brain regions. It is important with these infusions to distinguish between extinction and abolition via delayed diffusion to the essential region (Medina et al. 2002a; Zbarska et al. 2007, 2008). To do so, we began the infusion 40 min before the start of the session in one subject and 20 min before the session in the other subject. Since extinction requires fewer than 20 min, staggering the infusions in this way precludes the possibility that diffusion over 20 min abolished the conditioned responses (Medina et al. 2002a). As shown in Fig. 7, in both subjects the infusions induced extinction of delay $F_{(1,11)} = 17.127, P < 0.001$; blocks effect] and trace eyelid responses $F_{(1,11)} = 24.72, P < 0.001$; blocks effect] at rates similar to
those seen during ACSF infusion ($P = 0.73$ for delay; $P = 0.33$ for trace; blocks $\times$ session interaction). These data show that the cerebellar mechanisms known to operate for delay responses can support the extinction of trace responses under circumstances where noncerebellar mechanisms are not engaged.

We also used a second method to test whether cerebellar mechanisms are sufficient for extinction of trace responses when noncerebellar mechanisms are bypassed. This test involved presenting inputs that roughly match the duration of tone and PFC-driven inputs to cerebellum that occur during trace conditioning via electrical stimulation of mossy fibers through two separate electrodes (see Kalmbach et al. 2010a). The patterns of stimulation were designed to mimic the temporal duration of mossy fiber inputs to the cerebellum present during trace eyelid conditioning. Two stimulating electrodes were implanted in the middle cerebellar peduncle, where mossy fibers enter the cerebellum. A 500-ms train of pulses was delivered through one electrode to mimic tone-driven mossy fiber inputs, and a 1,000-ms train of pulses was delivered to the second electrode to mimic mossy fiber inputs driven by persistent activity neurons in mPFC (Aitkin and Boyd 1978; Boyd and Aitkin 1976; Campolattaro et al. 2011; Freeman and Muckler 2003; Siegel et al. 2009; Takehara-Nishiuchi and McNaughton 2008). We have found previously that responses acquired using these patterns of mossy fiber stimulation quantitatively reproduce the properties of trace conditioned responses (Kalmbach et al. 2010a). In this study we tested whether extinction of these trace-like responses is prevented when cerebellar-mediated extinction mechanisms are blocked via infusions of muscimol into the DCN. The experimental design was identical to that employed in the first experiment (e.g., Figs. 3 and 4).

In five subjects, the effects of inactivating the DCN on extinction were assessed using the same test sessions as presented in Fig. 2 following training with mossy fiber stimulation as the CS (example electrode placements are shown in Fig. 8E). As before, infusing muscimol into the DCN abolished the expression of conditioned responses [Fig. 8; $23.6 \pm 3.9\%$ responding during ACSF infusion, $2.2 \pm 2.2\%$ responding during muscimol infusion; $t_{(4)} = 4.13$, $P = 0.004$; paired $t$-test]. However, this time the muscimol infusion prevented the
extinction of stimulation-mediated trace-like responses. The average response likelihood during the entire postmuscimol session and the control 1 session was not different \((P = 0.14; \text{paired sample } t\text{-test})\), but both were different from response likelihood during the control 2 session \([\text{Fig. 8, B and C; } t_{(6)} = 4.09, P = 0.01 \text{ for postmuscimol; } t_{(9)} = 6.25, P = 0.003 \text{ for control 1; paired } t\text{-tests}]\). Furthermore, the rates at which responses extinguished during control 1 and postmuscimol sessions were not different \((\text{compare light gray and dark gray curves in Fig. 8B; } P = 0.67; \text{blocks } \times \text{session interaction})\). Figure 8 also shows the difference scores, which demonstrate that response likelihood during the control 1 and postmuscimol sessions did not differ across training blocks, whereas response likelihood during the control 2 session was less than that during postmuscimol sessions for the first couple of blocks. Together, these data provide evidence that the previously identified cerebellar mechanisms of extinction of delay responses can also mediate extinction of trace-like eyelid responses when noncerebellar mechanisms are bypassed by the procedures of this experiment. Thus extinction of trace eyelid responses can be mediated by different sites, one in the cerebellum and one outside of the cerebellum.

The rate of extinction for trace eyelid responses shows a bimodal distribution. Finally, we tested a second corollary of the hypothesis, stating that two different mechanisms can, in different subjects or under different circumstances, mediate extinction of trace responses. If, as hinted by the above-described data (Fig. 6), these mechanisms impose somewhat different rates of extinction, then we should expect a bimodal distribution of extinction rates to be apparent in a sufficiently large sample of data. To test this corollary, we examined the number of conditioned responses and the trials to criterion extinction \((\text{the first instance of } 8 \text{ non-conditioned responses in } 9 \text{ consecutive trials})\) in 21 trace-conditioned subjects. As shown in Fig. 9, the distribution of both parameters differed from an expected normal distribution \(\text{(black curve, Fig. 9A)}\), with the majority of subjects extinguishing within the first couple of blocks and a minority extinguishing more slowly, nearly halfway through the session \(\left(P < 0.01 \text{ for both trials to criterion and number of conditioned responses; Kolmogorov-Smirnov test for normality}\right)\). This apparent bimodality can be seen most clearly by comparing the trials to reach criterion with the number of conditioned responses during the extinction session for each subject \(\text{(Fig. 9B)}\). Also shown in Fig. 9 are behavioral sweeps from a subject that extinguished quickly and another that extinguished slowly. These data are consistent with the idea that the extinction of trace conditioned responses can be mediated by mechanisms operating in at least two sites and that extinction mediated by one site is faster than extinction mediated by the other site.

**DISCUSSION**

These results provide evidence that extinction of forebrain-dependent trace eyelid responses can occur via mechanisms located in more than one brain region. We used a within-animal design to compare the effects of blocking a signal that is necessary for extinction mechanisms in the cerebellum on delay and trace responses \(\text{(see Table 1 for a summary of the experiments and key comparisons)}\). Whether this signal was blocked directly by infusions of gabazine in the inferior olive \((\text{Medina et al. 2002a)}\) or indirectly by inactivation of the DCN \((\text{Hardiman et al. 1996; Ramnani and Yeo 1996)}\) \(\text{(Fig. 4)}\), we observed results consistent with previous findings: the extinction of delay responses was prevented. Our key finding is that these manipulations do not necessarily prevent the extinction of trace conditioned responses. Indeed, in most animals given inferior olive infusions and in all animals given DCN infusions, trace conditioned responses extinguished even while delay conditioned responses did not \(\text{(Fig. 6)}\). Apparent differences between DCN and inferior olive manipulations on trace extinction may be due to differences in the sensitivity of each test. For the muscimol in the DCN infusions, effects on extinction were inferred indirectly based on animals responding the next day, whereas the effects of gabazine into the olive could be seen directly during the first day of extinction. Nevertheless, because trace extinction was relatively unaffected by manipulations of two distinct regions separated by several millimeters, it is likely that decreases in trace eyelid response were caused by extinction rather than the abolition of responses \(\text{(e.g., by diffusion of the drug to a currently uniden-}

**Fig. 7.** Infusion of 2,3-dihydroxy-6-nitro-7-sulfamoyl-benzo(f)quinoxaline (NBQX) into the inferior olive causes extinction of delay and trace conditioned responses. Response likelihood for delay \((A)\) and trace conditioning trials \((B)\) are shown for the gabazine, NBQX, and ACSF infusion sessions. Eyelid position sweeps for the 2 subjects that received the NBQX infusion are shown in C. Symbols correspond to the data presented in Fig. 6. Notice that despite the fact that the CS and US were paired, learned responses diminished at a rate similar to that during extinction training while ACSF was infused. Eyelid closure in the lightest gray regions represents reflexive responses to the US.
tified region necessary for trace conditioning). Thus the data support the hypothesis that the extinction of delay conditioned responses is signaled by the inhibition of climbing fibers by collaterals in the DCN (Bengtsson et al., 2004; Hesslow and Ivarsson 1996; Medina et al. 2002a; Svensson et al. 2006), whereas trace conditioned responses can undergo extinction independent of this process.

The observation that for some subjects the gabazine infusions in the inferior olive prevented trace extinction (Fig. 6) suggests that the cerebellar mechanisms can support extinction of trace responses when, for whatever reasons, the noncerebellar mechanism is not engaged. In two separate experiments we tested the hypothesis that the cerebellum mechanism is sufficient to mediate extinction of trace responses when the noncerebellar mechanism is bypassed. First, we found that blocking the US input to the cerebellum (Mauk et al. 1986; McCormick et al. 1985; Medina et al. 2002a) with infusions into the inferior olive of the AMPA receptor antagonist NBQX caused the extinction of delay and trace conditioned responses (Fig. 7), even though each CS was being paired with the US. Consistent with a previous report (Medina et al. 2002a), this effect was independent of when (20 or 40 min) the infusion was delivered before the start of the session. Thus the decrease in responding after the infusion does not reflect the delayed extinction of the trace conditioned responses.
abolition of responses due to “cerebellar malfunction” as previously suggested (Zbarska et al. 2007, 2008) but is instead dependent on the number of training trials presented after the infusion. Second, using direct stimulation of mossy fibers to mimic the tone and mPFC-driven input to the cerebellum that occur during trace conditioning (Aitkin and Boyd 1978; Boyd and Aitkin 1976; Freeman and Muckler 2003; Kalmbach et al., 2010a; Takehara-Nishiuchi and McNaughton 2008), we provide evidence that extinction can occur through cerebellar mechanisms under circumstances where the noncerebellar mecha-

Table 1. Summary of experiments and key comparisons

<table>
<thead>
<tr>
<th>Experiment (Fig.)</th>
<th>Key Comparisons</th>
<th>Behavioral Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muscimol (1 mM) or ACSF control in deep nucleus during extinction (Figs. 2–4)</td>
<td>Postmuscimol day 1 responding vs. control day 1 (ACSF infusion) and control day 2 for both delay 500 and trace 500/500</td>
<td>If muscimol blocked extinction (delay or trace), then postmuscimol responding = control day 1, both &gt;control day 2</td>
</tr>
<tr>
<td>Gabazine (20 μM) in inferior olive during extinction (Fig. 5–6)</td>
<td>Responding during extinction (gabazine vs. ACSF) for both delay 500 and trace 500/500</td>
<td>If gabazine blocked extinction, then responding during gabazine infusion &gt; responding during ACSF infusion</td>
</tr>
<tr>
<td>NBQX (150 μM) during paired training (Fig. 7)</td>
<td>NBQX during paired conditioning vs. ACSF during extinction training for both delay 500 and trace 500/500</td>
<td>If NBQX caused extinction, then responding during ACSF extinction = responding during NBQX training</td>
</tr>
<tr>
<td>Muscimol (1 mM) in deep nucleus during extinction with mossy fiber stimulation as the CS (Fig. 8)</td>
<td>Postmuscimol day 1 vs. control day 1 (ACSF infusion) and control day 2</td>
<td>If extinction is blocked, then postmuscimol = control day 1, both &gt;control day 2</td>
</tr>
<tr>
<td>Rate of extinction during trace conditioning for 21 subjects trained with tone CS (Fig. 9)</td>
<td>Trace extinction: distribution of trials to criterion and number of CRs vs. predicted normal distribution</td>
<td>Cumulative frequency of trials to criterion and number of CRs diverge from expected normal distribution</td>
</tr>
</tbody>
</table>

ACSF, artificial cerebrospinal fluid; CR, conditioned response; CS, conditioned stimulus; NBQX, AMPA receptor antagonist 2,3-dihydroxy-6-nitro-7-sulfamoyl-benzo([f]quinoxaline.
that the rate of extinction for trace responses should display a bimodal distribution. Examining trace responses in 21 subjects, we found that the rate of extinction during the first day of extinction training was indeed bimodally distributed; most animals extinguished quickly within the first couple of blocks, whereas a few subjects extinguished more slowly approximately halfway through the session (Fig. 9). Although these data are consistent with the hypothesis that trace conditioned responses can extinguish via more than one mechanism and that these mechanisms produce different rates of extinction, other explanations are possible.

What is the noncerebellar mechanism of trace extinction? Although our experiments do not directly address the location or mechanism of the noncerebellar mechanism, there are at least two general possibilities: 1) descending brain systems could inhibit the motor systems that control eyelid closure, or 2) the mPFC-driven input to the cerebellum that is required for the expression of trace conditioned responses could diminish or be inhibited (Kalmbach et al. 2009).

In either case, the cerebellum-independent extinction process may occur in the forebrain. Whereas lesions or reversible inactivation of the cerebellum affect both delay and trace eyelid conditioning (Kalmbach et al. 2009; McCormick and Thompson 1984; Pakaprot et al. 2009; Woodruff-Pak et al. 1985), only trace conditioning is affected by lesions of forebrain structures such as the hippocampus, mPFC, and primary sensory cortex (Galvez et al. 2007; Kalmbach et al. 2009; Kim et al. 1995; Moyer et al. 1990; Powell et al. 2001; Solomon et al. 1986; Takehara et al. 2003; Weible et al. 2000). Thus it is possible that plasticity in these sites can contribute to the extinction of trace conditioned responses. Indeed, lesions of hippocampus, infralimbic area of mPFC, and other forebrain structures have been reported to affect the extinction of trace conditioned responses with long trace intervals, short trace intervals, and even delay conditioned responses (McCormick and Thompson 1982; Moyer et al. 1990; Weible et al. 2000). Ongoing studies are investigating the extracerebellar sites involved in the extinction of trace conditioned responses.

The first possible extracerebellar extinction mechanism, that descending brain systems inhibit the motor systems that control eyelid closure, parallels evidence that the extinction of learned fear is partly mediated by the inhibition of the brain systems responsible for the expression of fear memories (Maren and Quirk 2004; Milad and Quirk 2002; Milad et al. 2004; Quirk and Mueller 2008). In eyelid conditioning, learning in the cerebellum is expressed through its connections with the red nucleus, which in turn projects to the premotor and motor neurons that innervate the orbicularis oculi muscle responsible for eyelid closure (Morcuende et al. 2002; Thompson and Steinmetz 2009). Because these areas receive monosynaptic and polysynaptic input from various forebrain regions (Bernays et al. 1988; Morcuende et al. 2002), the noncerebellar extinction mechanism could parallel the fear extinction pathways by involving descending inhibition of these sites. The present work and data from previous studies, however, impose a specific constraint on this mechanism: this downstream inhibition would have to occur only for trace and not for delay conditioning. This constraint is imposed by observations that manipulations of cerebellum, but not red nucleus, prevent the extinction of delay conditioned responses (Figs. 3 and 4) (Hardiman et al. 1996; Medina et al. 2002a; Ramnani and Yeo 1996; Robleto and Thompson 2008).

The second possibility, that mPFC-driven input to the cerebellum diminishes or is inhibited, is based on recent evidence showing that learning in the cerebellum during trace eyelid conditioning is in response to mPFC-driven input that persists through the trace interval (Kalmbach et al. 2009, 2010a; Siegel et al. 2009). This is in contrast to delay conditioning, where cerebellar learning occurs in response to mossy fiber inputs that are more directly driven by the auditory system (Campolattaro et al. 2007; Freeman et al. 2007; Halverson et al. 2008; Halverson and Freeman 2010; Steinmetz et al. 1987). This difference suggests that extinction of trace conditioned responses could occur through a currently unidentified mechanism that results in the omission of mPFC-driven input to the cerebellum.

Although the plausibility of this hypothesis must be evaluated by future experiments, we propose the following concrete but speculative hypothesis regarding the mechanisms of extinction for trace eyelid responses. During the acquisition of trace eyelid conditioning, neurons in prefrontal cortex acquire the ability to respond to the CS with activity that persists through the trace interval to overlap with the US and in this way engage cerebellar learning. This input is necessary for the expression of trace eyelid responses as revealed by reversible abolition of trace response expression by inactivation of mPFC with infusion of the GABA_A receptor agonist muscimol (Kalmbach et al. 2009). As such, the noncerebellar mechanism of extinction may involve changes in this forebrain-driven response that render it unable to support expression of conditioned responses (either suppressing it or shortening its duration). This hypothesis is consistent with the relatively abrupt disappearance of trace responses during extinction: once the forebrain-driven response is omitted or foreshortened, trace responses would disappear immediately. To the cerebellum, it would be the equivalent of omitting the CS. We further hypothesize that, for unknown reasons, this forebrain mechanism is sometimes not successful, in which case the cerebellar mechanism would be sufficient to mediate extinction of the trace responses. This aspect of the hypothesis is consistent with the bimodal distribution of extinction rates we present in Fig. 9. It is also consistent with evidence that the extinction of trace eyelid responses can occur in the cerebellum when the required forebrain input is ensured (Figs. 7 and 8). Ongoing single-unit recording studies are testing this hypothesis more directly. These studies should reveal neurons in mPFC and pontine nuclei that show tone-evoked persistent activity that extinguishes at the same rate as behavioral extinction in some animals and more slowly than behavioral extinction in others.

In summary, we have provided evidence that trace eyelid responses can be extinguished via processes operating at more than one site. This shows that when the expression of a learned response requires plasticity in multiple brain systems, as is the case for trace conditioning, not only can extinction of that response be mediated by processes operating at more than one site, but there may be subject-to-subject or situation-to-situation differences in the relative contributions from those sites. Trace eyelid conditioning offers a unique opportunity to identify the variables that determine the factors, sites, and mechanisms by which such responses extinguish.

J Neurophysiol • doi:10.1152/jn.00381.2011 • www.jn.org
REFERENCES


