Behavioral/Systems/Cognitive

Eyelid Conditioning to a Target Amplitude: Adding How Much to Whether and When

Joy C. Kreider[1] and **Michael D. Mauk**[1,2]

[1]Center for Learning and Memory, [2]Section of Neurobiology, University of Texas at Austin, Austin, Texas 78712

Conceptual and practical advantages of pavlovian eyelid conditioning facilitate analysis of cerebellar computation and learning. Even so, eyelid conditioning procedures are unrealistic in an important way. The error signal to the olivocerebellar system does not decrease as learning adapts response amplitude or gain. This inherently limits the utility of eyelid conditioning for studies investigating how cerebellar learning mechanisms acquire and store an adaptive response amplitude. We report the development and characterization of a training procedure in which conditioned response amplitude is brought under experimental control with contingencies that more closely parallel natural conditions. In this procedure, the delivery of the unconditioned stimulus (US) is made contingent on conditioned response amplitude: the US is delivered for responses that fail to reach a specified target amplitude and is omitted for responses that meet or exceed the target. We find that rabbits trained with either a tone or with mossy fiber stimulation as the conditioned stimulus learn responses that approach target amplitudes ranging from 2 to 5 mm. Inactivating the interpositus nucleus with muscimol infusions abolished these conditioned responses, indicating that cerebellar involvement in eyelid conditioning is not tied explicitly to the use of pavlovian procedures. Together with previous studies, these data suggest that response amplitude is learned and encoded in the cerebellum during eyelid conditioning. As such, these results provide a foundation for systematic and controlled investigations of the cerebellar mechanisms that learn and encode the proper amplitude of adaptive movements.

Introduction

Pavlovian eyelid conditioning has facilitated analysis of cerebellar learning and computation (McCormick and Thompson, 1984; Perrett et al., 1993; Mauk and Donegan, 1997; Medina and Mauk, 1999; Hansel et al., 2001; Medina et al., 2001, 2002), mostly because of the close correspondence between the training stimuli and cerebellar inputs as well as between cerebellar output and behavioral responses (McCormick and Thompson, 1984; Mauk et al., 1986; Steinmetz et al., 1986; Mauk and Donegan, 1997; Hesslow et al., 1999). The notion, dating to Pavlov (1927), that pavlovian conditioning is "reflexive learning" in which the conditioned stimulus (CS) acquires through learning the ability to activate the reflex pathway triggered by the reinforcing stimulus is a misconception commonly applied to eyelid conditioning. Indeed, the misnomer "eyeblink" conditioning apparently stems from this misconception. Data contrary to this idea include the known circuitry of eyelid conditioning, which clearly show that the CS pathway is not connected to the reflex pathway through learning (Mauk and Donegan, 1997; Thompson et al., 1998), from evidence that conditioned eyelid responses are not blinks (Schade Powers et al., 2010), and from the simple behavioral observation that during acquisition conditioned response amplitude is not all-or-none. With training, conditioned response amplitude grows monotonically from barely detectable to maximum eyelid closure.

Even so, the standard training procedures of eyelid conditioning do not bring response amplitude under experimental control. For more natural responses such as adaption of the vestibulo-ocular reflex (VOR), the error input to the olivocerebellar system (i.e., to the climbing fiber input to the cerebellum) decreases as learning adapts response gain or amplitude (Gilbert and Thach, 1977; Raymond and Lisberger, 1997, 1998). Because the likelihood of the error signal to the olivocerebellar system [via presentation of the unconditioned stimulus (US)] does not decrease as conditioned response amplitude grows, pavlovian procedures inherently promote conditioned responses (CRs) at maximum amplitude (5–6 mm for rabbits). Even with air puff as the US, the CR (extension of the nictitating membrane) may not fully eliminate US activation of the climbing fibers because the external eyelids are generally restrained and the nictitating membrane itself is quite sensitive. For these reasons, eyelid conditioning has not been useful for the analysis of the cerebellar mechanisms involved in learning and encoding response amplitude (i.e., response gain).

We report the development and characterization of an eyelid training procedure that more closely mimics natural demands on cerebellar learning and that brings response amplitude under experimental control. In this procedure, US delivery is made contingent on conditioned response amplitude just as error inputs to the cerebellum depend on response gain. When response amplitude meets or exceeds a prespecified target amplitude, US delivery is omitted (see Fig. 1). We demonstrate that response amplitude approaches target amplitudes ranging from 2 to 5 mm,
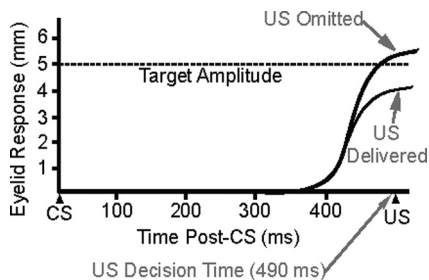
Figure 1. A schematic representation of the training procedures used to train eyelid responses to a target amplitude. Each training trial involved the presentation of a tone CS and the scheduled delivery of the eye stimulation US at an interstimulus interval of 500 ms. The actual delivery of the US was made contingent on response amplitude by monitoring the amplitude of the response in real time during the CS. Ten milliseconds before the scheduled delivery of the US (US Decision Time), the training software was programmed to determine whether the US would be delivered: for responses that had met or exceed the target amplitude, the US was delivered; for responses that had not achieved the target amplitude, the US was delivered.
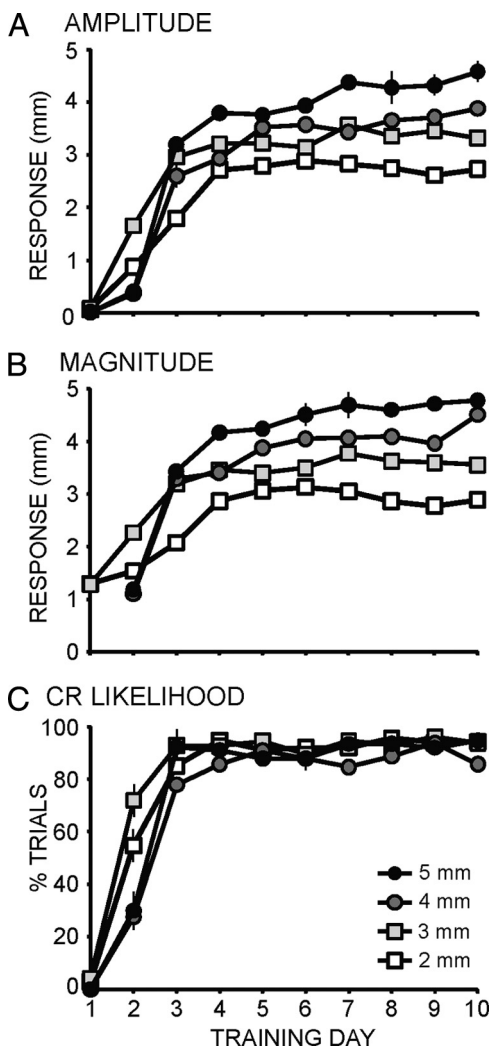

Figure 2. The acquisition of conditioned eyelid responses over 10 d of training for four groups that were assigned different target amplitudes (2, 3, 4, and 5 mm). **A**, The amplitude of the conditioned eyelid responses over the 10 d of training shows acquisition of different response amplitudes for the four groups. **B**, Similar results were obtained for the response magnitude measure, which excluded responses that did not achieve the 0.3 mm criterion to be considered a CR. **C**, Because the likelihood of responding was robust for all four groups, the results for amplitude and magnitude were quite similar. In instances in which error bars (±SEM) are not visible, the error is smaller than the symbol.
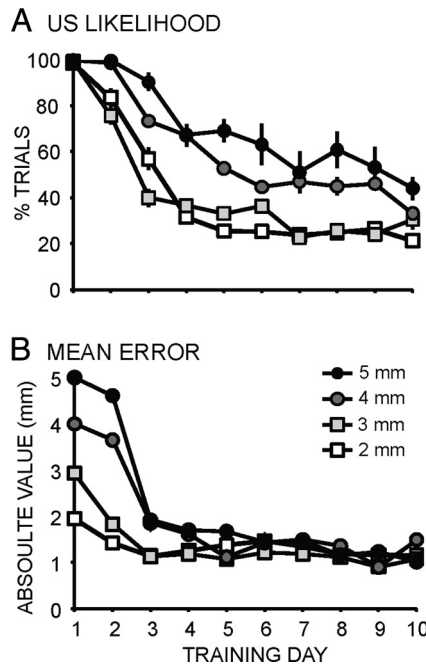

Figure 3. **A**, **B**, Decreases in the likelihood of US delivery (**A**) and in the mean error (**B**) as a function of 10 d of training for the four amplitude groups (2, 3, 4, and 5 mm). A decline in US likelihood (**A**) indicates an increase in the proportion of responses that met or exceeded the target amplitude. All four groups showed significant decreases in US likelihood, arriving at values ranging from 21 to 44%. **B**, The mean error decreased to similar values in all four groups.

with significant group differences present when analyzed after 10 training sessions. Cerebellar involvement in this learning task is supported by two observations: (1) response amplitude can be learned when mossy fiber stimulation substitutes for the tone, and (2) inactivating cerebellar output via infusion of muscimol into the interpositus nucleus abolishes expression of CRs, suggesting that cerebellar learning is responsible not only for whether and when to respond but also for how much to respond. These data indicate that this new training procedure represents a controlled experimental paradigm to study cerebellar mechanisms that learn and encode response amplitude.

Materials and Methods

Subjects. Thirty-two male albino rabbits (*Oryctolagus cunniculus*) weighing 2–3 kg were used. Animals were individually housed, were fed daily, and had access to water *ad libitum*. The care and treatment of these animals was in accordance with the standards of the Institutional Animal Care and Use Committee and National Institutes of Health guidelines.

Surgery. Animals were first anesthetized with a ketamine/acepromazine mixture and placed in a stereotaxic restrainer. Anesthesia was then maintained with isoflurane (1–4% mixed with oxygen), and sterile technique was used throughout the surgery. A 3 cm midline incision provided access to the skull. Each animal was prepared with anchor screws mounted in the skull and a head bolt that was cemented to the anchor screws with dental acrylic. For the eight animals involved in the mossy fiber stimulation and cerebellar infusion studies, the surgery included a craniotomy to permit the insertion of two tungsten stimulating electrodes (tip exposed to obtain ~100–200 kΩ; A-M Systems) that were aimed at the middle cerebellar peduncle and a 26 gauge stainless steel guide cannula (Plastics One) aimed at the interpositus nucleus ipsilateral to the training eye (all coordinates referenced to lambda; +3 mm anteroposterior, −5.5 mm mediolateral, −16 mm dorsoventral). Dental acrylic was added to affix these implants to the skull screws and head bolt. After the incision was closed, two stainless steel stimulating electrodes were implanted above and one just rostral and the other just caudal to the left
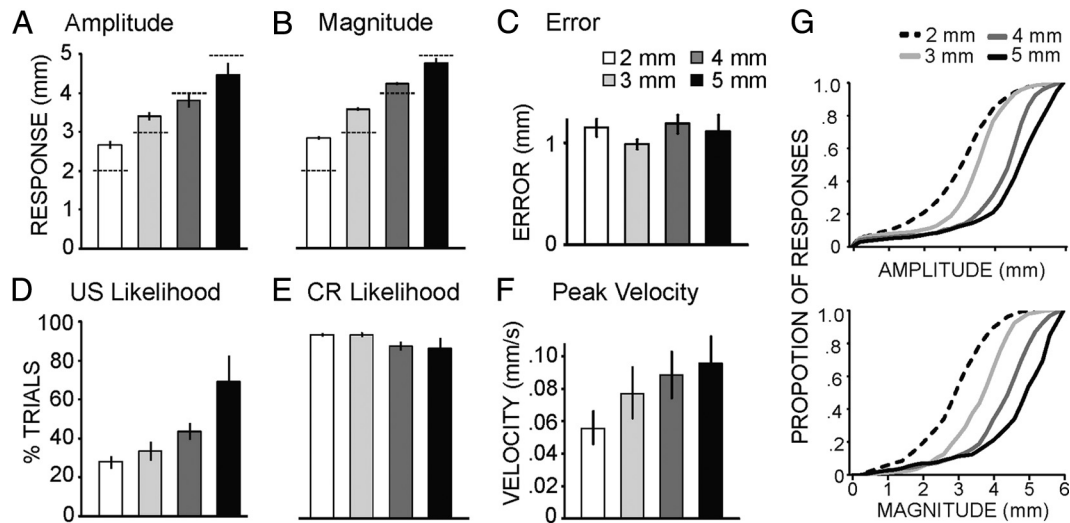
Figure 4. Comparisons of the four amplitude groups at steady-state performance taken for each animal from days 9 and 10. **A**, The steady-state amplitude of responding differed for the four groups. The four dotted lines indicate the target amplitude for each group. All pairwise comparisons (Tukey's test, $p < 0.05$) were significant, except for the comparison between the 3 and 4 mm group. **B**, Same format as **A** for magnitude data, which are the average amplitude of all CRs with the non-CRs excluded. All pairwise comparisons were significant (Tukey's test, $p < 0.05$). **C**, The steady-state average error (absolute value) did not differ significantly for the four groups. **D**, There was a general trend for the proportion of trials in which the US was delivered to increase with increased target amplitude. The only pairwise comparison that was statistically reliable was 2 versus 5 mm. **E**, All four groups show comparable performance in terms of CR likelihood; there were no significant differences between the four groups. **F**, There was a general trend for increased peak velocity of the CRs as the target amplitude increased. The only significant (Tukey's test, $p < 0.05$) comparison was 2 versus 5 mm. **G**, Cumulative proportion of responses as a function of conditioned response amplitude (top) and magnitude (bottom) for the four groups. Each CR for each animal was binned (0.2 mm bin width). All six pairwise comparisons are statistically reliable (Kolmogorov–Smirnov test, $p < 0.005$) for amplitude and also for the magnitude measure (except 4 vs 5 mm, $p < 0.05$).
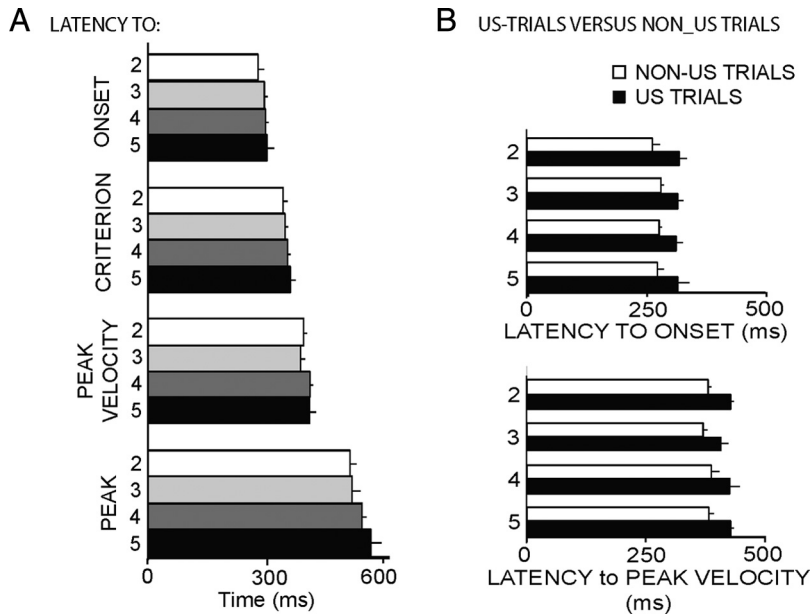

Figure 5. Measures of response timing compared across the four amplitude groups. **A**, The four graphs show latency to onset, latency to criterion latency to peak velocity, and latency to peak amplitude (for non-US trials). There were no statistically reliable difference between the groups for these response measures. **B**, Timing differences between US trials and non-US trials. Separate analysis of US trials (amplitude below target) and non-US trials (amplitude at or above target) revealed a general trend for faster, shorter-latency responses on non-US trials. All eight comparisons (4 groups × 2 response measures) were statistically reliable (paired t test, $p < 0.01$).

eye. Antibiotics and analgesics were administered after surgery. The rabbits recovered for at least 1 week before testing began.

Apparatus. Animals were trained in two custom-designed, electrically shielded chambers. The internal dimensions of each chamber were 90 × 61 × 61 cm (width × length × height). A wooden partition divided the chamber so that two rabbits could be trained simultaneously, each held in a standard plastic rabbit restrainer. The chambers were equipped with

a speaker connected to an audio source (model V85- 05; Coulbourn Instruments) that generated tones and a pair of isolated pulse stimulators (model 2100; A-M Systems) that delivered constant-current pulses through the periorbital electrodes. An infrared emitter detector was attached to the head bolt of each rabbit and directed at the left eye to measure eyelid position. For animals trained with mossy fiber stimulation as the CS, stimulus isolators (model A360; World Precision Instruments) coupled to an additional stimulator were used to deliver constant-current pulses through wires connected to the implanted electrodes via gold pins. Stimulus presentation and data collection were controlled by custom software. Eyelid position data were sampled at 1 kHz, and individual sweeps consisting of 2500 data points (200 ms before the CS provided a baseline and 2300 ms after CS onset) were collected for each trial and stored for subsequent offline analysis.

Training and testing procedures. The training procedures were a variation of standard delay eyelid conditioning. As with numerous previous studies, the CS was a 550 ms tone (1.0 kHz, 85 dB, 5 ms rise time). For the mossy fiber stimulation experiment, the CS was cathodal stimulation (100 Hz, 40 μs pulse width, 100 μA). The US was a 50 ms pulse train (100 Hz, 1 ms pulse width, ~1–2.5 mA constant current) delivered through the periorbital electrodes and presented 500 ms after CS onset so that the US and CS ended together.

The US intensity was adjusted for each rabbit to a level that induced robust reflex responses but did not elicit pain reactions. The infrared eyelid position detector was calibrated eliciting a full reflex response and defining the amplitude of that response as 6 mm.

To attempt to train responses to a specific amplitude, US delivery was made contingent on response amplitude relative to a preselected target amplitude (Fig. 1). During each trial, 10 ms before the scheduled delivery of the US, the maximum response amplitude to that point was compared with the target amplitude for that animal. When the response amplitude met or exceeded the target amplitude for at least four 1 ms bins, the US was omitted; otherwise, the US was presented. Each session consisted of eight blocks of nine trials with an intertrial interval of 30 ± 10 s. All animals were trained for 10 d. Using these procedures, four groups of animals were trained, each with a different target amplitude (2, 3, 4, and 5 mm).

Drug infusions. Eight of the 32 rabbits (four trained using the tone CS and four using mossy fiber stimulation) were included in the infusion experiment. These animals underwent 10 d of acquisition training and then on subsequent training days received drug infusions or vehicle infusions. The GABA$_A$ agonist muscimol (Tocris Bioscience) was dissolved at a concentration of 1 mM in the sterile artificial CSF vehicle consisting of the following (in mM): 119 NaCl, 2.5 KCl, 1.2 NaH$_2$PO$_4$, 2 MgCl$_2$, 2 CaCl$_2$, 26 NaHCO$_3$, 10 D-glucose, and 20 HEPES, pH 7.35–7.4. The solution was delivered through a 33 gauge internal cannula that extended 1.2, 1.5, or 1.7 mm beyond the guide cannula (if infusions were ineffective at 1.2 mm, longer cannula were used). The internal cannula attached to polyethylene PE 50 tubing was coupled to a 50 μl Hamilton syringe that was mounted on an automated injector system (model MD-1001; Bioanalytical Systems) that was driven by an electric pump (model MD-1020). The order of vehicle and drug infusions was counterbalanced. Training was paused after the fourth block, the cannula was inserted, and the infusion began. A total volume of 2 μl was infused over 20 min, and then training resumed. At least 1 d of retraining was given between infusion sessions to ensure there were no long-lasting effects.

Data analysis. Offline computer analysis using custom software was conducted on the digitized eyelid position data stored from each session. The following measures of eyelid closure were determined for each trial on each training day. First, eyelid position measured 200 ms before each trial established a baseline for detecting movement during each trial. A CR was defined as an eyelid closure of at least 0.3 mm with an onset between 40 ms and 1 s after CS onset (analysis terminated at US offset for US-delivered trials). For trials that included a CR, analysis also included latency to the 0.3 mm criterion, latency to onset as determined by an algorithm that detects initial deflection from baseline, peak velocity, latency to peak velocity, and latency to peak amplitude. Two separate measures of response size were used. Amplitude refers to the average amplitude for all trials, including those in which there was no CRs (i.e., amplitude < 0.3 mms). Magnitude refers to the average amplitude for only those trials in which there was a CR. For all measures, trials were excluded from analysis if a movement greater than 0.3 mm occurred during the 200 ms baseline period before CS onset.

Statistics. Tests for differences were conducted using the Kolmogorov–Smirnov test, one-way ANOVA, two way ANOVA mixed design, two-way ANOVA with repeated measures, and paired and unpaired t tests. Post-ANOVA comparisons were assessed using the Tukey's test (α was $p < 0.05$). Two animals were removed from the 5 mm group as a result of technical difficulties (error in calibration).

Histology. Histology was performed for animals that had mossy fiber stimulating electrodes and infusion cannula. Stimulation sites were marked by passing direct anodal current (200 μA for 20 s) through a stainless steel wire that was cut to the length of the internal cannula and guided through the guide cannula or to the stimulating electrode, respectively. Animals were killed with an overdose of sodium pentobarbital and were perfused through the heart with 0.9% saline and then with 10% Formalin. Brains were embedded in gelatin, mounted on a freezing stage, and sectioned at 80 μm using a sliding microtome. Sections were mounted and stained with cresyl violet.

Results

Acquisition of a target response amplitude

Under natural conditions, such as adaptation of the VOR, pursuit, or saccades, error-driven excitatory drive onto the climbing fiber input to the cerebellum decreases as response gain ap-
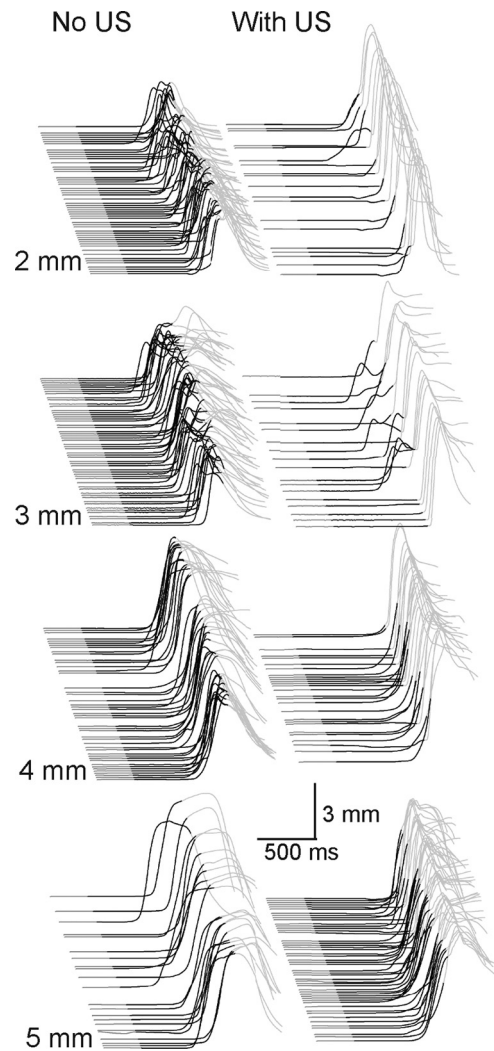

Figure 6. Sample raw responses from four animals, one from each amplitude group, taken from the day 10 of training. Each sweep is a single eyelid response (first trial in front), with the black region of each sweep indicating the duration of the CS. The responses were divided into non-US trials (left column) and US trials (right column). The sequence of US and non-US trials for each animal is apparent in the spacing of the trials. Note the general trend for more non-US trials at small target amplitudes and more US trials for larger target amplitudes. The reflex response component of each trial is apparent in the light gray region of the "With-US" trials.

proaches the correct value (Gilbert and Thach, 1977; Raymond and Lisberger, 1997, 1998). Our training procedure is intended to more closely approximate these natural learning conditions by coupling US presentation to response amplitude relative to a target amplitude. To do so, we trained four groups of animals for 10 d, with each group subjected to one of four target amplitudes (2, 3, 4, or 5 mm). Twenty-four rabbits were randomly assigned to one of these groups, with two animals from the 5 mm group eventually excluded for technical difficulties. For each daily session, the US was omitted on trials in which the eyelid response reached the target amplitude within 10 ms of the time of US delivery. The US was delivered on trials in which the eyelid response failed to reach the target amplitude.

As shown in Figure 2, animals in the four groups acquired robust CRs, and, for each group, the average amplitude/magnitude of the responses approach the assigned target. For the amplitude measure, all responses were included, whereas the magnitude measure excluded all non-CRs (those with amplitude < 0.3 mm). Separate two-way mixed ANOVAs (one for ampli-
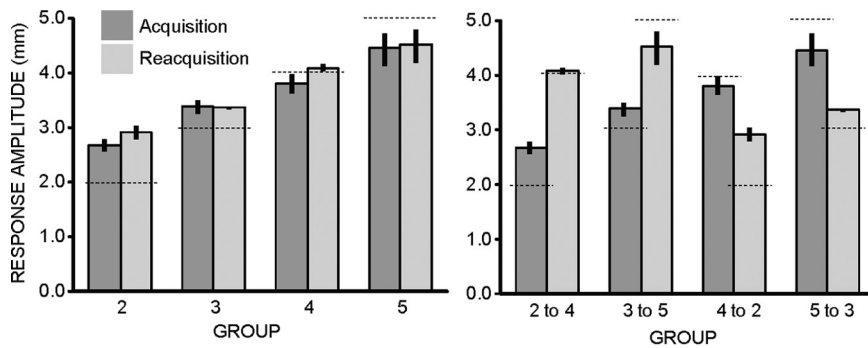
Figure 7. A comparison of steady-state response amplitudes after initial acquisition (dark gray; same data as Fig. 2A) and in the same animals after 10 d of reacquisition to a new target amplitude. The initial target amplitude and reacquisition target amplitudes were, respectively, for the four groups (2 to 4 mm, 3 to 5 mm, 4 to 2 mm, and 5 to 3 mm). As in Figure 4, the data are derived from an average for each animal over acquisition (and reacquisition) days 9 and 10. Left, A comparison of identical target amplitudes for acquisition and reacquisition. There were no statistically reliable differences for each target amplitude between the acquisition and reacquisition data (t tests). The target amplitude is indicated by the dotted lines. Right, A comparison of acquisition and reacquisition data for the same groups of animals. The different acquisition and reacquisition target amplitudes for each group are again indicated by dotted lines. Paired t tests ($p < 0.01$) indicated statistically reliable differences between acquisition and reacquisition amplitudes for each group.

tude and one for magnitude) each showed significant group (amplitude, $F_{(3,198)} = 7.43$, $p = 0.0039$; magnitude, $F = 11.18$, $p = 0.00013$), training day (amplitude, $F_{(9,162)} = 126.4$, $p < 0.0001$; magnitude, $F = 101.21$, $p < 0.0001$), and group × day interaction (amplitude, $F_{(27,162)} = 2.99$, $p < 0.0001$; magnitude, $F = 3.66$, $p < 0.0001$) effects. These analyses indicate that the animals learned, that there were differences between groups in the amplitude and magnitude of responses, and that these differences developed with training. Despite these differences in response amplitude and magnitude, the groups were comparable in terms of the proportion of trials in which a CR occurred (Fig. 2C). For these data, the two-way mixed ANOVA revealed a significant effect of training day ($F_{(9,162)} = 123.8$, $p < 0.0001$), but there were no group or group × training day interaction effects ($p > 0.05$).

There are two additional indicators that would support the hypothesis that this procedure brings response amplitude under experimental control and more closely parallels natural conditions. First, there should be a decrease in the proportion of trials in which the US is delivered from near 1.0 to an intermediate level, indicating that response amplitude is sometimes greater than and sometimes less than the target. Figure 3A shows that indeed the proportion of trials in which the US was delivered decreased over training, indicating an increase in the number of trials in which the target amplitude was met or exceeded. For this measure, there was a significant effect of groups ($F_{(3,198)} = 7.13$, $p = 0.0047$) and of training day ($F_{(9,162)} = 70.93$, $p < 0.0001$), but the interaction comparison was not statistically reliable ($F_{(27,162)} = 1.67$, $p = 0.0545$). Second, there should be a decline in the mean error, that is, the mean absolute value of the difference between response amplitude and the target amplitude. This measure is plotted as a function of training days in Figure 3B. An overall decrease in mean error is indicated by a significant effect of training days ($F_{(9,162)} = 102.17$, $p < 0.0001$). Because the initial error value differs for the four groups, it is not surprising that there was also a significant group effect ($F_{(3,198)} = 14.41$, $p < 0.0001$) and a significant group × training days effect ($F_{(27,162)} = 13.6$, $p < 0.0001$).

Overall, these analyses reveal that (1) the groups developed with training differences in response amplitude and magnitude, (2) the proportion of trials in which the US was delivered and the mean error both decreased over training, and (3) there are no

differences in likelihood of responding for the three groups. Because the most straightforward way to assess the properties of this learning is to evaluate performance at or near steady state, we next focused analysis on data from the last 2 d of training.

Analysis of asymptotic performance
For analysis of steady-state performance, the data were pooled from acquisition days 9 and 10 for each subject. These analyses revealed clear differences in the amplitude and magnitude of the CRs between the four groups (Fig. 4A,B). Separate one-way ANOVAs for amplitude ($F_{(3,18)} = 19.53$, $p < 0.0001$) and magnitude ($F = 37.25$, $p < 0.0001$) support the reliability of these differences. For the magnitude measure, all six pairwise comparisons (Tukey's test, $p < 0.05$) were reliable, and for amplitude, only the 3 versus 4 mm comparison was not reliable. In contrast to these differences, there were no reliable differences in the mean error measure across the four groups (Fig. 3C). As a final test for the differences in amplitude and magnitude of the responses across the four target amplitude groups, normalized cumulative frequency distributions were calculated for each response measure (Fig. 4G). A Kolmogorov–Smirnov test ($p < 0.005$) revealed that all six pairwise comparisons for each group were statistically reliable.

Figure 4A–C combined suggest that, although the average amplitude error was comparable across the four groups, the source of the errors were different. Responses for the 2 and 3 mm groups tended to overshoot the target, whereas responses for the 5 mm group then tended to undershoot the target. This trend can be inferred from the comparison of US likelihood shown in Figure 4D, which reflects the proportion of responses that undershot the target. Despite the clear trend for this measure to increase with target amplitude, the differences between the groups only approached significance ($F_{(3,18)} = 3.44$, $p = 0.08$). Only the 2 mm group showed a significant tendency to overshoot the target because separate t tests for each group revealed only a reliable difference from a population mean of 50% (equal undershoot and overshoot responses) for the 2 mm group ($t_{(5)} = 3.47$, $p < 0.05$). Despite the difference in response amplitude and magnitude, the groups showed comparable performance ($F_{(3,18)} = 1.799$, $p = 0.37$) in terms of CR likelihood (Fig. 4E). Finally, Figure 4F shows that there was also a general trend for the peak velocity of the CRs to increase with the target amplitude ($F_{(3,18)} = 7.81$, $p = 0.003$). Pairwise comparisons showed that the 2 versus 4 mm and 2 versus 5 mm differences were reliable (Tukey's test, $p < 0.05$).

Previous studies have revealed that precise timing of the CRs is a characteristic feature of eyelid conditioning (Smith, 1968; Smith et al., 1969; Millenson et al., 1977; Mauk and Ruiz, 1992). The latencies to onset and rise times of the responses vary systematically with the interstimulus interval so that the responses tend to peak near the time of US delivery. We examined various measures of response timing for the four target amplitude groups and found no reliable differences in response timing (Fig. 5A). Neither the latency to onset, latency to criterion, latency to peak velocity, nor latency to response peak measures showed statisti-
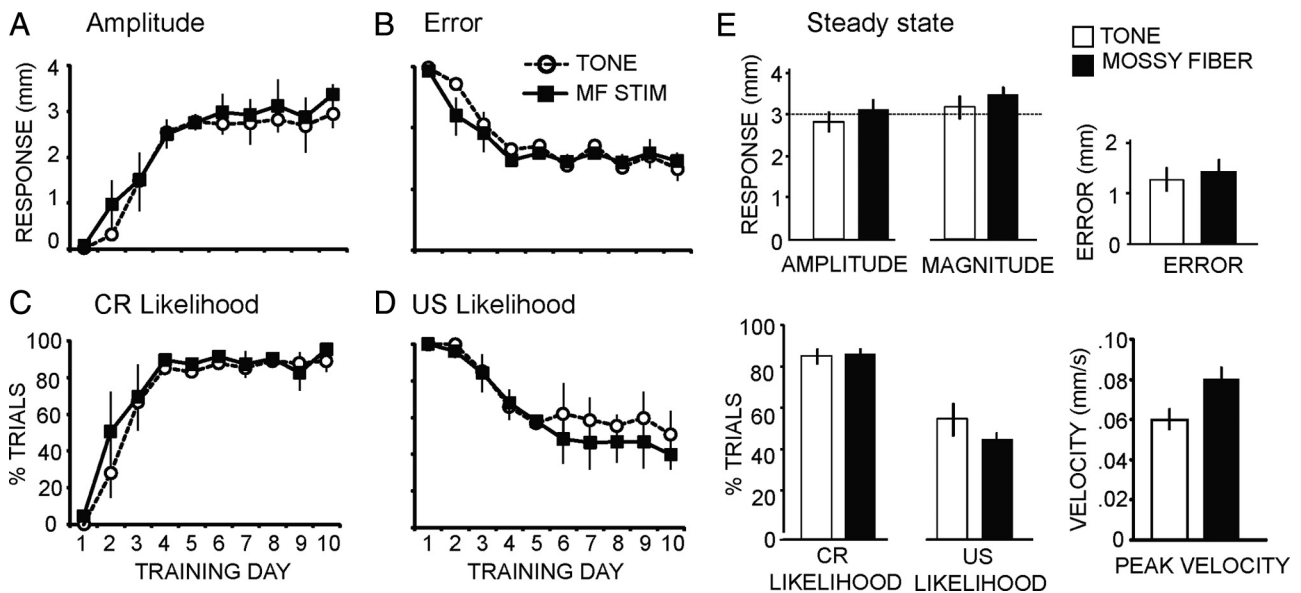
Figure 8. Acquisition to a target amplitude (3 mm) for animals trained with a tone CS and with mossy fiber stimulation as the CS. **A**, Amplitude of CRs over 10 d of training. **B**, Mean error (absolute value) over 10 d of training. MF, Mossy fiber. **C**, **D**, CR likelihood (**C**) and US likelihood (**D**) over the 10 d of training. **E**, Comparisons between tone CS and mossy fiber stimulation CS groups for various response measures taken from days 9 and 10 of training. None of the six group comparisons was statistically reliable (*t* test, *p* > 0.05).

cally reliable differences across groups. When responses were separated according to whether they overshot the target (non-US trials) or undershot the target (US trials), there were significant differences in the timing of the responses across all four groups (Fig. 5*B*). Separate two-way mixed ANOVAs for latency to onset and latency to peak velocity both showed only a significant effect of response type (non-US vs US; onset, $F_{(1,18)} = 100.6$, $p < 0.0001$; peak velocity, $F = 49.04$, $p < 0.0001$). This analysis demonstrates that in general the responses that met or exceeded the target amplitude had shorter latencies than did the responses that did not meet the target amplitude.

Sample raw responses from one subject from each amplitude group (day 10 in each case) are shown in Figure 6. For each subject, the responses that met or exceeded the target amplitude are shown in the left column, and those that did not meet target (and thus where the US was delivered) are shown in the right column. These data show the general trend for responses to overshoot for the smaller target amplitudes and to undershoot for the 5 mm target. Together, these analyses demonstrate that the target amplitude training procedure can be used to gain reasonable experimental control over response amplitude in eyelid conditioning, that the mean amplitude error are comparable across targets from 2 to 5 mm, and that these differences occur against comparable CR rates and comparable statistics for the timing of the CRs.

Response amplitude after reacquisition to a new target
Both the practical and conceptual importance of target amplitude training would be diminished if target response amplitudes could not be acquired and reacquired as dictated by the experimental design. We therefore tested each animal using a new target amplitude that was different from the initial target. After the 10 d of training to an initial target amplitude, each subject was then trained for 10 additional days using a new target. The new target amplitudes for the 2, 3, 4, and 5 mm groups, respectively, were 4, 5, 2, and 3 mm. Figure 7 shows a comparison of response amplitudes for these four groups for the initial and the new target amplitudes. As above, these data involved a pooled average of

days 9 and 10 for each subject. We began by asking whether there were differences between acquisition and reacquisition for each of the four target amplitudes (Fig. 7, left graph). This between-subject comparison compares each target when it was used in acquisition versus reacquisition (i.e., acquisition from the 2 mm group and reacquisition at the 2 mm target, originally the 4 mm group). A two-way ANOVA revealed only a main effect of target amplitude ($F_{(3,36)} = 37.06$, $p < 0.0001$), the acquisition versus reacquisition and interaction effects were not significant, and there were no reliable differences in the amplitudes obtained during acquisition versus reacquisition to a new target. The right portion of Figure 7 shows the same data replotted to highlight comparisons of the amplitudes obtained for acquisition and reacquisition in the same groups. Four separate paired *t* tests (*p* < 0.01) revealed that, for each group, the amplitude of the responses produced by acquisition to one target were different from those produced by reacquisition to a new target. These analyses indicate that both acquisition to a first target amplitude and reacquisition to a new target amplitude produced significant difference in response amplitude for each target amplitude group and that there were no differences in these trends between acquisition and reacquisition.

Mossy fiber stimulation as the CS
Considerable evidence indicates that the acquisition and expression of CRs during pavlovian (delay) eyelid conditioning is mediated by plasticity within the cerebellum. Because the conditioning to a target amplitude is not, strictly speaking, pavlovian, we conducted two additional experiments to test the possibility that learning under these circumstances is mediated by noncerebellar mechanisms. The ability to substitute stimulation of mossy fibers for the tone CS is among the stronger evidence supporting cerebellar mechanisms as the basis of eyelid conditioning. As such, we compared the effects of training to a target amplitude of 3 mm using a tone CS in one group and stimulation of mossy fibers as the CS in the second group. The procedures were otherwise identical to those used above for the 3 mm group. Figure 8 shows the results for 10 days of training for these two

groups. A–D compares these two groups over 10 d of training, and E shows the steady-state comparisons as performed in the previous experiment. Separate two-way mixed ANOVAs for the four response measures showed only main effects of training days (amplitude, $F_{(9,54)} = 30.54$; error, $F = 17.71$; CR likelihood, $F = 29.09$; US likelihood, 16.62; $p < 0.0001$ in all cases). None of the group main effects or interaction effects were significant, nor were the tone versus mossy fiber comparisons (Fig. 8 E) for the eight responses measures that were analyzed for the steady-state data (days 9 and 10, t tests, $p > 0.05$). These data exclude factors upstream of the cerebellum as contributing to the acquisition of response amplitude and support the hypothesis that eyelid conditioning to a target amplitude is mediated by the cerebellum.

Reversible inactivation of cerebellum

As an additional test of the hypothesis that the cerebellum is necessary for conditioning to a target amplitude, the expression of CRs was tested in the four tone-trained and four mossy fiber-trained animals before and after inactivation of cerebellar output via infusions of the GABA agonist muscimol into the anterior interpositus nucleus (Fig. 9). Subjects in both groups were tested on separate days with infusions of the vehicle or infusions of muscimol (order counterbalanced for each group). The group average graph in Figure 9 (middle) illustrates that, for both groups, there was no reliable change after vehicle infusion, whereas infusions of muscimol abolished the expression of the CRs. Raw data from one mossy fiber stimulation subject for the vehicle (left) and muscimol (right) infusions are shown at the top of Figure 9. Example cannula (left) and stimulating electrodes from this same subject are shown at bottom.

Together, the mossy fiber stimulation and deep nucleus inactivation studies show (1) that the cerebellum is necessary for the expression of CRs produced by eyelid conditioning to a target amplitude (infusion results), and (2) that the mechanisms responsible for acquiring and expressing the target response amplitude are not in the CR pathway upstream of the cerebellum (mossy fiber stimulation results).

Discussion

We have adapted traditional pavlovian eyelid conditioning procedures to make the error input driven by the US contingent on the amplitude of the CR. By omitting the US on trials in which the CR meets or exceeds a predefined target amplitude (and presenting the US otherwise), we have shown that the resulting conditioned eyelid responses tend toward the target amplitude. This is true for both acquisition to an initial target amplitude and reacquisition using a different target. Although there was a tendency for overshoot with small target amplitudes and for undershoot with the largest target, the mean absolute error did not differ for groups trained for target amplitudes ranging from 2 to 5 mm. These data show that, by adapting eyelid conditioning procedures to make better parallels with natural cerebellar learning circumstances, the amplitude of the CRs can be brought under experimental control. The ability to abolish these CRs with inactivation of the cerebellum and the ability to substitute stimulation of mossy fibers for the tone CS both support the hypothesis that, like pavlovian eyelid conditioning, eyelid conditioning to a target amplitude is mediated by the cerebellum. These results form the foundation for systematic and controlled studies of how the cerebellum learns, stores, and expresses the proper amplitude of the responses that it adapts.
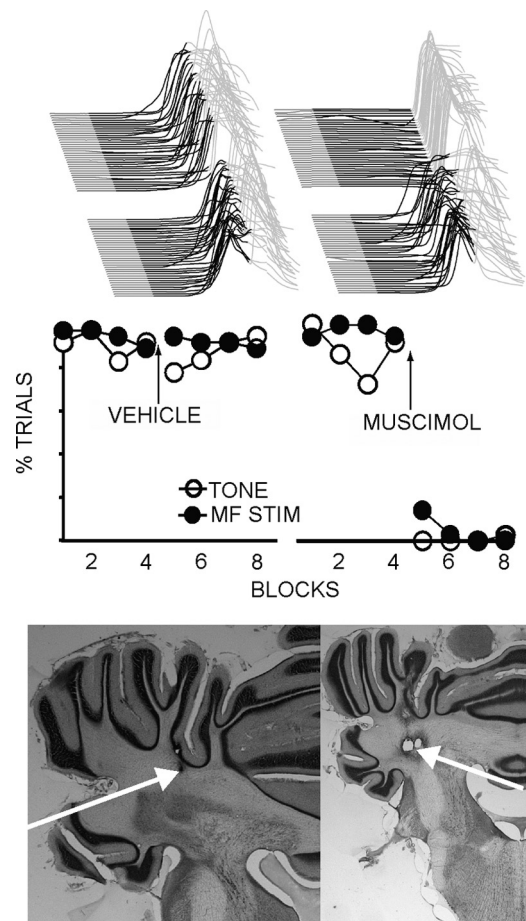

Figure 9. Necessity of the cerebellum for expression of responses acquired to a target amplitude of 3 mm. Two groups, one in which a tone CS was used and one in which fiber stimulation was used as the CS, were tested on separate days with either a vehicle infusion or with an infusion of the GABA agonist muscimol. The two graphs show group averages of the lack of effect of vehicle infusion (left) and the abolition of CRs with muscimol infusion (right). The time of infusion in each case is indicated with a black arrow. Sample raw data from one animal are shown above: traces on left are vehicle infusion, and the traces on the right are muscimol infusion. In each case, the preinfusion responses are in front (first trial in front), and the postinfusion responses are in the back. The two bottom photographs show placements of the infusion cannula (left) and the mossy fiber stimulating electrodes (right) for the animal whose raw data are shown above. Two mossy fiber electrodes were implanted in this animal; only one was used for this experiment. MF, Mossy fibers.

Implications for eyelid conditioning and the cerebellum

The necessity of the cerebellum for pavlovian delay eyelid conditioning is supported by a wealth of data. The present findings indicate that pavlovian contingencies, in which the delivery of the US does not depend on the CRs, is not a feature that makes the cerebellum necessary. Because, strictly speaking, procedures for training to a target amplitude involve operant contingencies, the necessity of the cerebellum in eyelid conditioning is related to the nature of the stimuli and the response system and not to the nature of the contingencies that govern the presentation of the reinforcing stimulus.

The present data also address a common misconception about eyelid conditioning, namely, that it is reflex conditioning or eyeblink conditioning. Pavlov initiated the idea that classical conditioning involves pathways activated by the CS acquiring the ability to activate the pathway that mediates the reflex response (Pavlov, 1927). The neural pathways known to mediate expression of the CRs and the reflex blinks elicited by the US are not

consistent with Pavlov's hypothesis. Reflex blinks elicited by primary afferents coursing through the superior orbital nerve are mediated by two pathways (Pellegrini et al., 1995). Primary afferents project to both the caudalis–interpolaris border region of the spinal trigeminal nucleus and to the dorsal horn, the C1 segment of spinal cord. Both of these regions project to the orbicularis oculi region of the facial nucleus. Lesion experiments have shown that the spinal trigeminal pathway mediates the initial R1 component of reflex blinks and the spinal cord pathway mediates the longer latency R2 component (Pelligrini et al., 1995). The CR pathway projecting from the anterior interpositus nucleus via the red nucleus appears to connect to this pathway at the motor or premotor neurons of the facial nucleus. Reflex blinks and CRs therefore are mediated by entirely different neural pathways and can be expected to display quite different properties, as has been established clearly in human studies (Schade Powers et al., 2010). Conditioned eyelid responses are therefore more properly characterized as "eyelid responses" than as "eyeblinks," and the present data demonstrate that the amplitude of these eyelid responses, like most other skeletal muscle responses, can be learned and mediated by cerebellar learning.

Implications for cerebellar learning of gain or amplitude

The question of how and where the cerebellum learns and stores response gain or amplitude is interrelated with the question of the site(s) of plasticity responsible for cerebellar learning. This question has been best addressed with studies involving analysis of adaptation of the VOR (Raymond et al., 1996; Boyden et al., 2004). Initial efforts in this area gave rise to considerable debates as to whether plasticity occurs in the cerebellar cortex or in the vestibular nucleus, whose connectivity and circuitry parallels deep cerebellar nuclei (Lisberger and Sejnowski, 1992, 1993; Ito, 1993). More recent theories suggest the involvement of plasticity in both cerebellar cortex and deep nucleus (Perrett et al., 1993; Raymond et al., 1996; Mauk and Donegan, 1997). With more than one site of plasticity involved, the question of the relative contributions of the two sites becomes important. The relative contributions of cerebellar cortex and cerebellar deep nucleus plasticity mechanisms to response gain or amplitude there remain unresolved, and there are indications that the relative contributions may vary in complex ways depending on the circumstances (Mauk, 1997). Eyelid conditioning to a target amplitude provides a way to address these questions with relatively simple and experimentally tractable system that is amenable to parametric behavioral analyses, reversible lesions, *in vivo* recordings, and even large-scale computer simulations.

References

Boyden ES, Katoh A, Raymond JL (2004) Cerebellum-dependent learning: the role of multiple plasticity mechanisms. Annu Rev Neurosci 27: 581–609.

Gilbert PF, Thach WT (1977) Purkinje cell activity during motor learning. Brain Res 128:309–328.

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

Hesslow G, Svensson P, Ivarsson M (1999) Learned movements elicited by direct stimulation of cerebellar mossy fiber afferents. Neuron 24: 179–185.

Ito M (1993) Cerebellar flocculus hypothesis. Nature 363:24–25.

Lisberger SG, Sejnowski TJ (1992) Motor learning in a recurrent network model based on the vestibulo-ocular reflex. Nature 360:159–161.

Lisberger SG, Morris EJ, Tychsen L (1987) Visual motion processing and sensori-motor integration for smooth pursuit eye movements. Annu Rev Neurosci 10:97–129.

Lisberger SG, Sejnowski TJ (1993) Reply to cerebellar flocculus hypothesis. Nature 363:25.

Mauk MD (1997) Relative contributions of cerebellar cortex and nuclei in motor learning: contradictions or clues? Neuron 18:343–346.

Mauk MD, Donegan NH (1997) A model of pavlovain eyelid conditioning based on synaptic organization of the cerebellum. Learn Mem 4:130–158.

Mauk MD, Ruiz BP (1992) The learning-dependent timing of Pavlovian eyelid responses: differential conditioning using multiple inter-stimulus intervals. Behav Neurosci 106:666–681.

Mauk MD, Steinmetz JE, Thompson RF (1986) Classical conditioning using stimulation of the inferior olive as the unconditioned stimulus. Proc Natl Acad Sci U S A 83:5349–5353.

McCormick DA, Thompson RF (1984) Cerebellum: essential involvement in the classically conditioned eyelid response. Science 223:296–299.

Medina JF, Mauk MD (1999) Simulations of cerebellar motor learning: computational analysis of plasticity at the mossy fiber to deep nucleus synapse. J Neurosci 19:7140–7151.

Medina JF, Garcia KS, Mauk MD (2001) A mechanism for savings in the cerebellum. J Neurosci 21:4081–4089.

Medina JF, Nores WL, Mauk MD (2002) Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. Nature 416: 330–333.

Millenson JR, Kehoe EJ, Gormezano I (1977) Classical conditioning of the rabbit's nictitating membrane response under fixed and mixed CS-US intervals. Learn Motiv 8:351–366.

Pavlov IP (1927) Conditioned reflexes: an investigation of the physiological activity of the cerebellar cortex. London: Oxford UP.

Pellegrini JJ, Horn AK, Evinger C (1995) The trigeminally evoked blink reflex. I. Neuronal circuits. Exp Brain Res 107:166–180.

Perrett SP, Ruiz BP, Mauk MD (1993) Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses. J Neurosci 13:1708–1718.

Schade Powers A, Coburn-Litvak P, Evinger C (2010) Conditioned eyelid movement is not a blink. J Neurophysiol 103:641–647.

Raymond JL, Lisberger SG (1997) Multiple subclasses of Purkinje cells in the primate floccular complex provide similar signals to guide learning in the vestibulo-ocular reflex. Learn Mem 3:503–518.

Raymond JL, Lisberger SG (1998) Neural learning rules for the vestibulo-ocular reflex. J Neurosci 18:9112–9129.

Raymond JL, Lisberger SG, Mauk MD (1996) The cerebellum: a neuronal learning machine? Science 272:1126–1131.

Smith MC (1968) CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response. J Comp Physiol Psychol 66:679–687.

Smith MC, Coleman SR, Gormezano I (1969) Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward CS-US intervals. J Comp Physiol Psychol 69:226–231.

Steinmetz JE, Rosen DJ, Chapman PF, Lavond DG, Thompson RF (1986) Classical conditioning of the rabbit eyelid response with a mossy fiber stimulation CS. I. Pontine nuclei and middle cerebellar peduncle stimulation. Behav Neurosci 100:878–887.

Thompson RF, Thompson JK, Kim JJ, Krupa DJ, Shinkman PG (1998) The nature of reinforcement in cerebellar learning. Neurobiol Learn Mem 70:150–176.