

# A Cholinergic Mechanism for Reward Timing within Primary Visual Cortex

Alexander A. Chubykin,<sup>1,3</sup> Emma B. Roach,<sup>2,3</sup> Mark F. Bear,<sup>1,\*</sup> and Marshall G. Hussain Shuler<sup>2,\*</sup>

<sup>1</sup>Howard Hughes Medical Institute, The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>2</sup>Department of Neuroscience, The Johns Hopkins University, Baltimore, MD 21205, USA

<sup>3</sup>These authors contributed equally to this work

\*Correspondence: [mbear@mit.edu](mailto:mbear@mit.edu) (M.F.B.), [shuler@jhmi.edu](mailto:shuler@jhmi.edu) (M.G.H.S.)

<http://dx.doi.org/10.1016/j.neuron.2012.12.039>

## SUMMARY

Neurons in rodent primary visual cortex (V1) relate operantly conditioned stimulus-reward intervals with modulated patterns of spiking output, but little is known about the locus or mechanism of this plasticity. Here we show that cholinergic basal forebrain projections to V1 are necessary for the neural acquisition, but not the expression, of reward timing in the visual cortex of awake, behaving animals. We then mimic reward timing in vitro by pairing white matter stimulation with muscarinic receptor activation at a fixed interval and show that this protocol results in the prolongation of electrically evoked spike train durations out to the conditioned interval. Together, these data suggest that V1 possesses the circuitry and plasticity to support reward time prediction learning and the cholinergic system serves as an important reinforcement signal which, in vivo, conveys to the cortex the outcome of behavior.

## INTRODUCTION

Predicting the timing and identity of salient events from limited sensory experience is critical for survival, allowing animals to exploit resources and avoid harmful situations. Learned temporal relationships are thought to be represented in the frontal and parietal cortices (Genovesio et al., 2009; Leon and Shadlen, 2003; Mita et al., 2009), regions atop a cortical hierarchy that receive highly processed sensory information (Felleman and Van Essen, 1991). Surprisingly, neurons within the primary visual cortex (V1), located at the lowest level of this presumptive hierarchy, are also capable of providing information about the learned timing of reward in relation to sensory input. This “reward timing activity” arises as a consequence of training adult rodents to associate visual cues with water reward at brief delays (e.g., 1 or 2 s). Through the course of conditioning, V1 neural responses evolve from relating simple features of the visual cues to expressing what these cues have come to predict: the expected time of reward (Shuler and Bear, 2006).

The means by which cortical neurons come to express reward timing activity are unknown but are thought to derive from a process of reinforcement learning, wherein a reinforcement signal relates the outcome of behavior with preceding neural activity (Dayan and Niv, 2008; Sutton and Barto, 1998). Computational studies have shown how such a reinforcement signal can impinge locally on synapses that have been active in the recent past, selectively modifying them so that appropriate cue-reward intervals are expressed (Gavornik and Shouval, 2011; Gavornik et al., 2009). These models assert that V1 receives convergent information about the stimulus as well as a signal conveying the subsequent receipt of reward. While it is well known that V1 receives feedforward visual input from the thalamus (Hubel and Wiesel, 1962), the source and identity of the hypothesized reinforcing signal are unknown, as are the mechanisms, perhaps local to V1, that transform the sensory response to encode time.

Neuromodulators are attractive candidates for reinforcement due to their ability to simultaneously broadcast a message of behavioral importance throughout the cortex (Doya, 2002; Woolf, 1996; but see Pennartz, 1995), and acetylcholine (ACh) from the basal forebrain (BF) is particularly well suited to reinforce V1 for a variety of reasons. There is a high density of cholinergic varicosities within V1 (Lysakowski et al., 1989; Mechawar et al., 2000) arising from multiple BF cholinergic nuclei, including the substantia innominata, nucleus basalis, and the diagonal band of Broca (Carey and Rieck, 1987; Rye et al., 1984). Functionally, ACh contributes to plasticity in V1 (Bear and Singer, 1986; Gu and Singer, 1993; for reviews see Gu, 2003; Origlia et al., 2008) and is involved in the alteration of tuning properties and map organization in other areas of cortex (Conner et al., 2003; Froemke et al., 2007; Kilgard and Merzenich, 1998; Sachdev et al., 1998; Weinberger, 2003). In addition, single-unit recordings in the BF have provided evidence for the encoding of unexpected rewarding (Santos-Benitez et al., 1995; Wilson and Rolls, 1990) and salient (Lin and Nicolelis, 2008) events.

The current study had two goals. The first was to determine whether the cholinergic innervation of visual cortex provides the conjectured reinforcement signal required for neurons to learn conditioned cue-reward intervals. The second was to investigate the possibility that V1, isolated from the rest of the brain in an in vitro preparation, could support response duration plasticity (RDP) analogous to reward timing activity observed in vivo.

## RESULTS

### Cholinergic Innervation Is Required to Learn, but Not Express, Reward Timing in V1

The first set of experiments utilized the behavioral protocol previously established in adult male Long-Evans rats by Shuler and Bear (2006) and was designed to distinguish between a requirement for local BF cholinergic input in the acquisition versus the expression of reward timing activity. Water-restricted rats were trained to approach a nose poke, receive one of two visual cues (100 ms, full-field retinal cue presented to either the left or right eye through removable, head-mounted goggles), and lick on a spout to receive delayed water reward (Figure 1A). Each cue was experimentally associated with a discrete number of licks necessary to release a drop of water. To correctly complete a trial, animals were required to lick  $m$  times after a presentation of cue 1 and  $n$  times after cue 2, translating to a short delay to reward after cue 1 and a long delay after cue 2. To differentiate reward receipt from expectancy, half the total number of trials were unrewarded even if the animal licked to criterion. Rats quickly adopt a simple strategy in this task: they begin licking immediately after cue presentation and exit the nose poke after either receiving and consuming a reward (rewarded trial, water delivery indicated by blue droplets in Figure 1A) or soon after the longer delay has expired (unrewarded trial, reward omission indicated by ghosted droplets in Figure 1A). Therefore, a feature of this task is that differences in neural activity across experimental conditions can be interpreted free of behavioral confound, as the exit times between the two unrewarded conditions are indistinguishable (see quantification below).

After the acquisition of stereotypical task performance (i.e., stable cue-reward intervals and nose poke exit times), animals were bilaterally implanted with microelectrode arrays and infusion cannulae in the deep layers of binocular V1. Single-unit data were collected while animals performed daily sessions of the task under the initial cue-reward pairing, in which cue 1 predicts a short delay to reward and cue 2 predicts a long delay. As previously observed (Shuler and Bear, 2006), many neurons appeared to report one (but not both) of the operantly conditioned intervals between visual cue and expected reward delivery with a spiking pattern of sustained increased, sustained decreased, or delayed peak activity (Figures 2A–2C).

To quantify this observation in individual neurons, we first used a correlation analysis to determine which cue evoked the greatest modulation of spiking activity, referred to as the neuron's "dominant" cue (as opposed to its "nondominant" cue) in the poststimulus period (>100 ms after stimulus offset up to the average trial duration). Second, we employed a bin-by-bin receiver operating characteristic (ROC) analysis to compare the responses evoked across all trials of dominant versus nondominant cue presentation. As exit behavior is consistent across trials containing presentations of either cue, this approach allows us to use the nondominant response as a comparison for any modulation resulting from the animal's immediate behavior. Of 842 well-isolated single units, 609 (48% sustained increase, 36% sustained decrease, and 16% peak) showed dominant cue-evoked spiking activity in the post-

stimulus period that differed significantly from the nondominant response and were considered for further analysis.

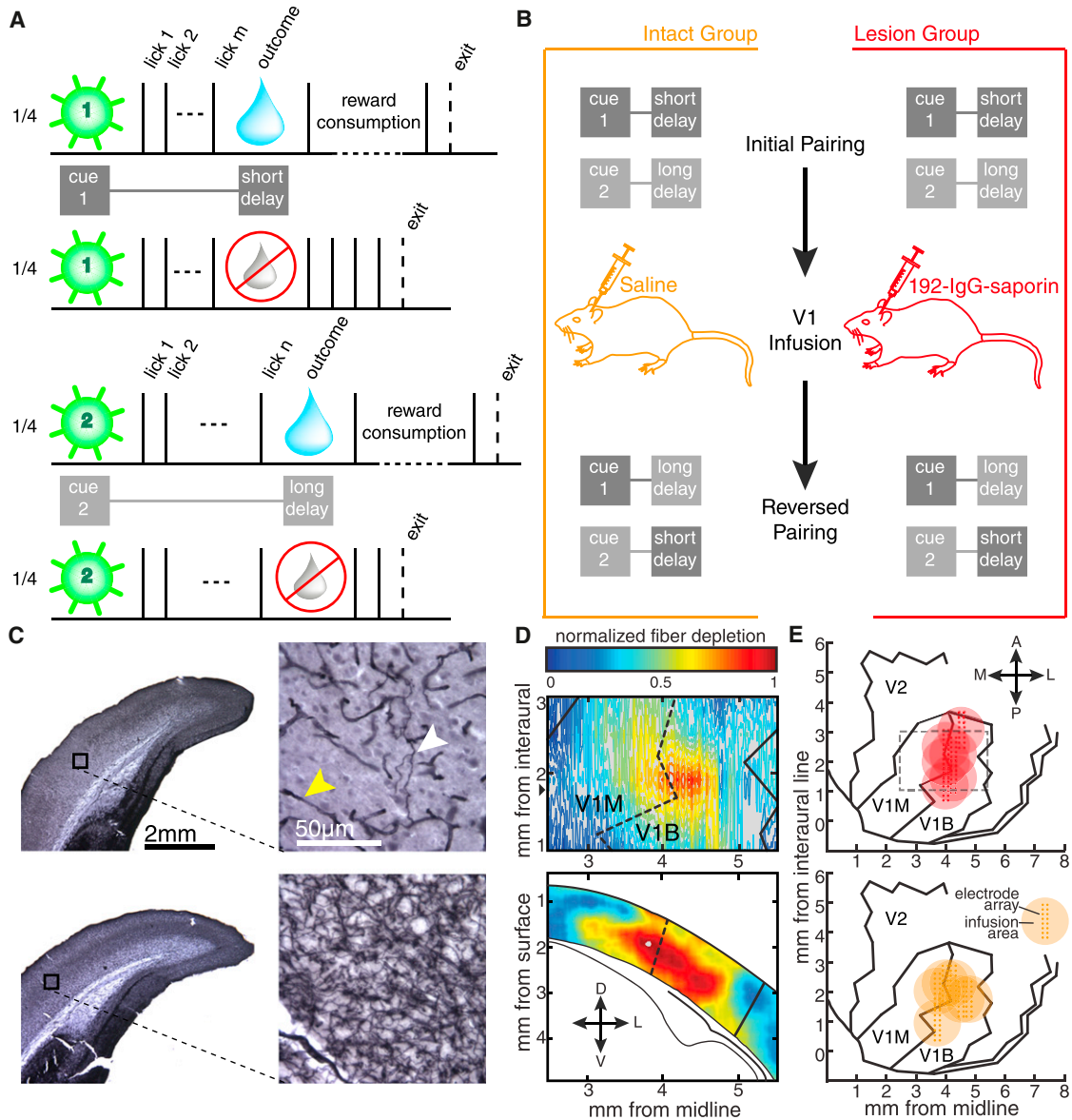
For each of these units, the neural report of time (NRT) was defined as the moment the area under the ROC curve returned to chance with 95% confidence or, for peak neurons, as the poststimulus moment with the highest ROC value (Figure 2C). To analyze NRTs collected across animals that experienced different cue-reward intervals, we expressed NRTs by a "delay index" score:

$$\text{delay index} = \frac{\text{NRT} - \tau_{\text{short}}}{\tau_{\text{long}} - \tau_{\text{short}}},$$

where  $\tau_{\text{short}}$  and  $\tau_{\text{long}}$  correspond to the average time the animal received reward after the short and long delay, respectively. Under this index, a value of zero translates to a perfect report of the short delay, whereas a value of one is a perfect report of the long delay. If neurons tend to report the expected reward time associated with their dominant cue, then a population analysis of cue 1-dominant neurons should reveal that NRT scores conform to the short delay and that cue 2-dominant neurons match the long delay. The cumulative distributions of NRT scores shown in Figure 2D demonstrate that, indeed, the subpopulations have significantly different median values that accord with the operantly conditioned reward delays paired with their dominant cues. These results corroborate the original *in vivo* description of cue-specific reward timing activity in V1.

Having established that subpopulations of V1 neurons exhibit reward timing to the delay paired with their dominant cue, we went on to query the necessity of BF cholinergic innervation in (1) the neural expression of the previously learned contingency and (2) the ability for neurons to acquire reward timing that matches novel task parameters. To address these issues, animals first received a bilateral infusion of either 192-IgG-saporin ( $n = 6$ ) or saline ( $n = 7$ ) through cannulae that were implanted alongside and centered to each recording electrode array (Figure 1E) and were then subjected to additional training with reversed cue-reward pairings (Figure 1B). The immunotoxin 192-IgG-saporin binds p75, a growth factor receptor expressed exclusively by cholinergic neurons arising from the BF, and is transported to the soma where it induces cell death by ribosome inactivation (Wiley et al., 1991). Importantly, the technique is fast (Waite et al., 1994) and eliminates only cholinergic BF input to the infusion zone, leaving the rest of the circuit intact (Holley et al., 1994). Based on measures of dye diffusion in the rat brain (Myers, 1966), and given our cannula/electrode geometry, we expected our infusion (~350 nl) to impact a 1- to 2-mm-diameter spherical volume centered about the electrode array. We used a diaminobenzidine-enhanced Hatchett's Brown stain for acetylcholinesterase (Tago et al., 1986) (Figure 1C) to visualize the expanse of our lesioning protocol, both in area and in depth (Figure 1D), and quantified the extent of lesion as a volume with a diameter of  $1.583 \pm 0.105$  mm (mean  $\pm$  SE,  $n = 3$ ; see Supplemental Experimental Procedures available online). The fiber loss was restricted to the occipital cortex and was not observed in any subcortical region.

The selective removal of cholinergic input to V1 prior to task reversal enabled us to disambiguate the effect of cholinergic depletion upon the expression of previously established



**Figure 1. Experimental Design**

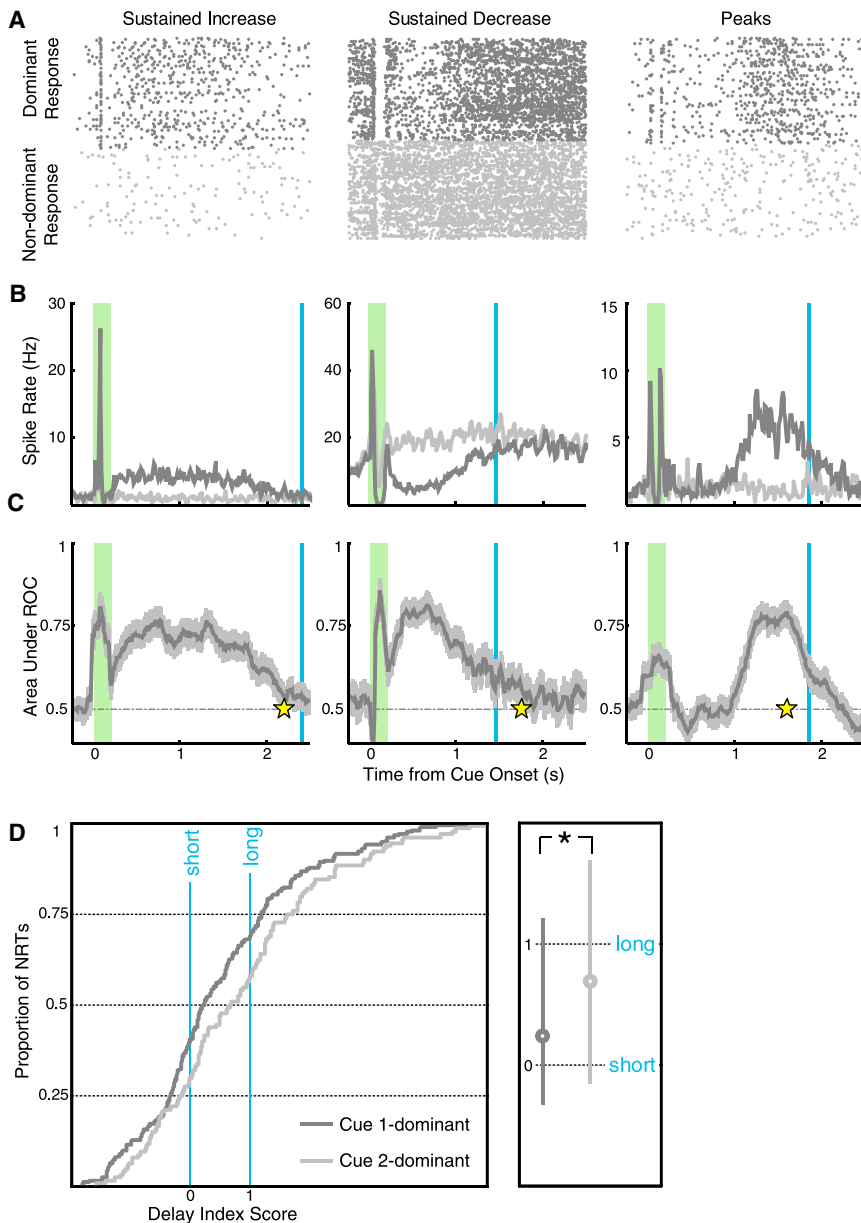
(A) After a 2 s intertrial interval, rats could enter the nose poke to activate one of four pseudorandomly interleaved trial types. A brief cue (green flash) was presented to either the left or the right eye and was predictive of the number of licks (solid black ticks) required to gain a small bolus of water (blue drop) in half of the trials.

(B) Implanted animals performed the task as outlined in (A) before receiving a bilateral infusion of either saline or 192-IgG-saporin into V1. After recovery, the task parameters were reversed such that the cue previously associated with the short delay was now paired with the longer delay and vice versa.

(C) Coronal sections demonstrating AChE histochemistry from a lesioned animal (top) and an intact animal (bottom). The black boxes on the low-magnification images (left) correspond to the high-magnification regions (right). The white arrowhead indicates a remaining fiber stained for AChE, and the yellow arrowhead indicates a capillary.

(D) Example visualization of fiber depletion. The contour plot (top) shows the relative difference of laminar-averaged staining intensity between atlas-matched sections from an intact and lesioned hemisphere. The heat plot (bottom) represents the relative comparison for the coronal slice with the widest apparent lesion extent (anterior-posterior position indicated by the black triangle next to the contour plot above). The solid lines demarcate the boundary of V1, while the dashed lines indicate the border between the monocular (V1M) and binocular (V1B) portions of V1. Bottom diagram is adapted from Paxinos and Watson (2008).

(E) Aerial view of approximate recording locations and infusion zones for the lesion (top) and intact (bottom) groups. Each large circle indicates the expected region affected by the infusion for each animal, and the estimated position of the 8 × 2 recording electrode array is indicated by the pin points. The gray dashed line indicates the region shown in the contour plot in (D). Aerial maps are adapted from Zilles (1985).



**Figure 2. V1 Neurons Recorded under Initial Pairing**

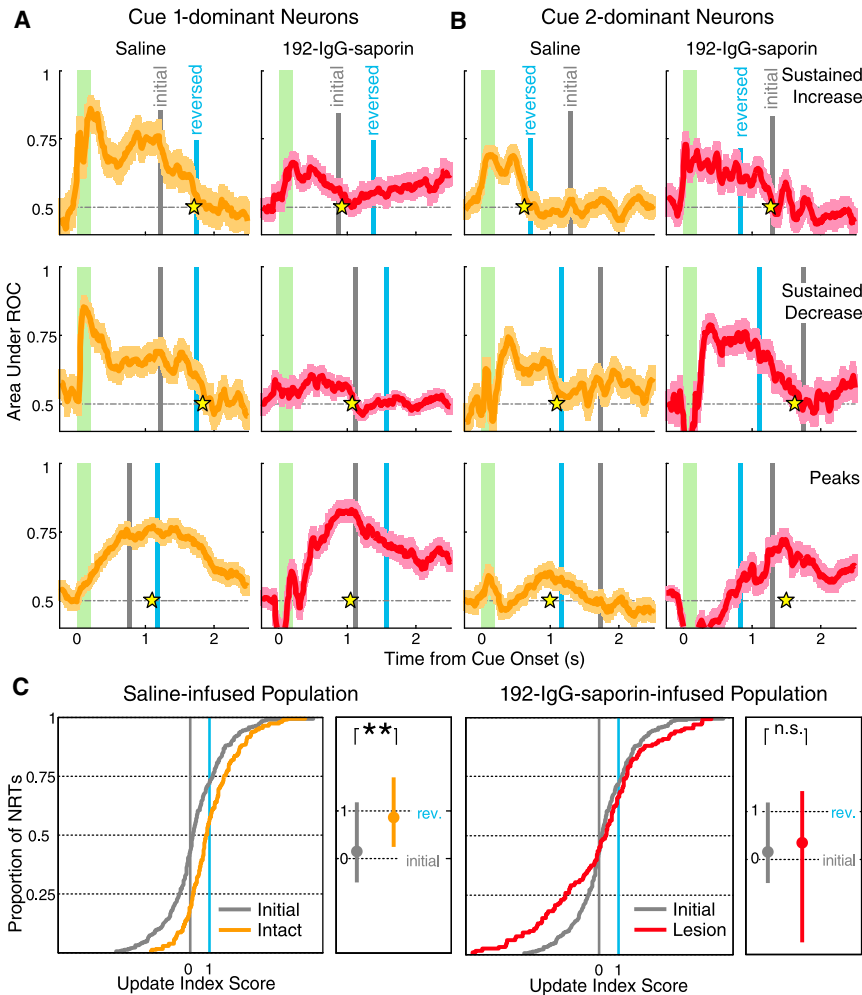
(A–C) Columns contain single-unit examples of each form of reward timing (left, sustained increase; center, sustained decrease; right, peaks). (A) Dots represent spikes recorded from a single unit in a behavioral session, with responses to the dominant cue on top (dark gray) and nondominant below (light gray). The raster plots are gathered from all correctly completed unrewarded trials, aligned at stimulus onset, and stacked in chronological order. (B) Smoothed, average spike rates compiled across trials are shown in dark gray for the dominant response and light gray for the non-dominant (shaded green bar: cue presentation; blue line: average reward delivery after the dominant cue). (C) Bin-by-bin ROC analysis, comparing responses evoked in all trials of dominant and nondominant cue presentation, provides the area under each ROC curve (dark gray) and a 95% confidence interval (light gray). For sustained increase (left) and sustained decrease (center) neurons, the neural report of time (NRT; yellow star) is defined as the first bin returning to chance. For peak neurons (right), the neural report of time is defined as the maximal ROC value.

(D) Cumulative population distributions of delay index scores are plotted on the left, where zero represents the short delay to reward and one represents the long delay. Plotted in dark gray is the population of neurons ( $n = 180$ ) whose response is dominated by the cue associated with the short delay (cue 1) and in light gray the population ( $n = 116$ ) dominated by the cue paired with the longer delay (cue 2). The median values of these distributions lie close to zero and one, respectively, and are significantly different from one another (Mann-Whitney U test,  $p < 0.05$ ), indicating that the observed subpopulations accurately relate the reward times associated with their dominant cues. The boxplot on the right indicates the median values (circle) and 25%–75% percentile range (line) for each subpopulation.

reward timing versus the acquisition of new reward timing intervals. Therefore, after 5–6 days of postinfusion recovery, single-unit data were collected while animals performed daily sessions of the task with reversed parameters (i.e., cue 1 predicting a long delay and cue 2 predicting a short delay; Figure 1B). Because animals exit the nose poke shortly after the expiration of the longer delay time on unrewarded trials, regardless of which cue was presented (Student's paired  $t$  test,  $p = 0.1862$ ), the behavioral readout (average unrewarded exit time divided by average long reward delay) remained stable after cue-delay reversal for both intact and lesioned rats (two-way repeated-measures ANOVA: task epoch,  $p = 0.744$ ; infusion content,  $p = 0.179$ ; interaction,  $p = 0.105$ ). The average delays to reward were also unchanged across reversal

task epoch,  $p = 0.653$ ; infusion content,  $p = 0.246$ ; interaction,  $p = 0.155$ ).

Figures 3A and 3B show example neurons of each form of reward timing (sustained increase, sustained decrease, and peak) collected after the reversal for neurons that were either cue 1 dominant (Figure 3A) or cue 2 dominant (Figure 3B). These examples illustrate that both cue 1-dominant and cue 2-dominant neurons recorded from intact animals express NRTs that appropriately reflect the new policy. Conversely, although cue 1- and cue 2-dominant neurons recorded from 192-IgG-saporin-infused animals are capable of displaying all forms of reward timing activity, they do not update their NRTs but rather persist in reporting the now outdated policy.



**Figure 3. V1 Neurons Recorded after Cortical Infusion and Cue-Reward Reversal**

(A and B) Example neurons exhibiting sustained increase (top), sustained decrease (middle), and peak (bottom) reward timing recorded from intact (orange) and lesioned (red) animals after V1 drug infusion and cue-reward pairing reversal. Conventions are as in Figure 2, with the current reward time associated with the dominant cue shown in blue (“reversed”) and the reward time initially associated with the dominant cue shown in gray (“initial”). (A) Cue 1-dominant neurons from saline-infused animals (left) update their policy to reflect the reversed contingency (cue 1 associated with a long delay), while neurons from lesioned animals (right) continue to report the initial policy. (B) Cue 2-dominant neurons report the new contingency (cue 2 associated with a short delay) in intact animals (left), while neurons from 192-IgG-saporin-infused animals (right) continue to express the outdated policy.

(C) NRTs recorded after contingency reversal, plotted as cumulative population distributions of update index scores. Zero represents the reward time initially associated with the dominant cue and one represents the reward time associated with the dominant cue after reversal (gray, observations recorded under the initial task parameters; orange, observations collected from intact animals after reversal; red, observations gathered from lesioned animals after reversal). Scores from intact animals (left;  $n = 182$ ) form a population with a median value that is distinct from the initial observations (Mann-Whitney U test,  $p < 10^{-10}$ ), while neurons collected from lesioned animals (right;  $n = 131$ ) continue as a population to report the same median value ( $p = 0.6929$ ). See also Figures S1 and S2. Boxplot conventions are as in Figure 2.

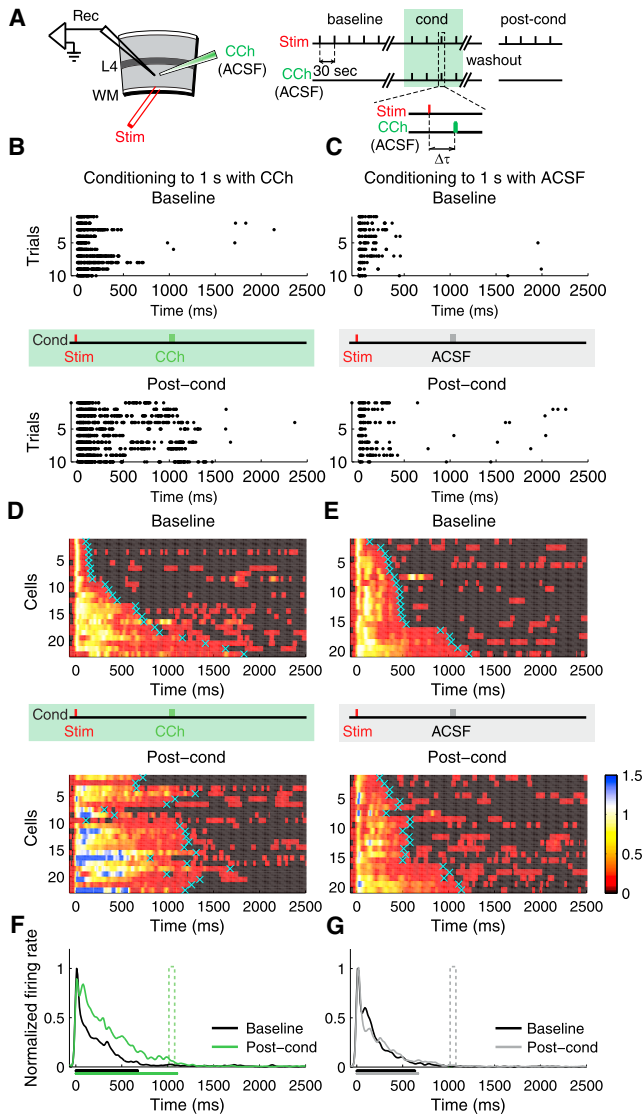
To quantify these observations from 192-IgG-saporin-infused and intact animals after policy reversal, we scored NRTs gathered under the second contingency using an “update index” to express the degree to which NRTs reflected the new, as opposed to the previously conditioned, reward time. The update index expresses the temporal difference of the NRT from the previously associated time of reward in units of the absolute difference between the old and new reward times:

$$update\ index = \begin{cases} cue\ 1-dominant, \tau_{reversed} > \tau_{initial} : \frac{NRT - \tau_{initial}}{|\tau_{initial} - \tau_{reversed}|} \\ cue\ 2-dominant, \tau_{reversed} < \tau_{initial} : \frac{\tau_{initial} - NRT}{|\tau_{initial} - \tau_{reversed}|} \end{cases}$$

where  $\tau_{initial}$  and  $\tau_{reversed}$  correspond to the average reward delays paired with the dominant cue during the initial and reversed task parameters. Under this “update index,” a score of zero indicates that the NRT has not moved from the initially associated reward time, whereas a positive one indicates traveling exactly one unit toward the new reward time (a negative one indicates that the NRT has moved a complete unit in the wrong direction). Prereversal NRTs were also expressed by

this index for the purpose of comparing population responses before and after policy reversal. Since prereversal index scores from intact and 192-IgG-saporin-infused animals were not significantly different from each other (cue 1-dominant: Mann-Whitney U test,  $p = 0.4010$ ; cue 2-dominant:  $p = 0.7980$ ), these data were combined together for the remaining analyses.

The distribution of update index scores collected from intact, saline-infused animals has a median value that is not significantly different from a score of one (Wilcoxon signed-rank test,  $p = 0.8161$ ) and shows a significant shift relative to the prereversal data (Figure 3C, left), indicating that as a population, neurons from intact animals updated their policy to reflect the novel pairing. However, the median value of the neural population recorded from 192-IgG-saporin-infused animals postreversal is not significantly different from that observed under the initial conditions (Figure 3C, right). Indeed, the median value is indistinguishable from the initial target index (Wilcoxon signed-rank test,  $p = 0.8219$ ), indicating that the population is continuing to express the previously learned temporal relationships. Dividing the neurons by cue dominance into subpopulations reveals the same effect: neuronal responses from intact animals come to reflect the newly paired cue-reward intervals whether their



**Figure 4. Response Duration Plasticity**

(A) Experimental design (Rec, recording electrode; Stim, stimulating electrode; WM, white matter; L4, layer 4; CCh, carbachol; ACSF, artificial cerebrospinal fluid). Conditioning (cond) is performed by CCh application at  $\Delta t$  delay after electrical stimulation.

(B and C) Raster plots of representative neurons before (baseline) and after (post-cond) conditioning. (B) A neuron conditioned with CCh applied at 1 s delay. (C) A control neuron conditioned with ACSF applied at 1 s delay. The vertical red tick indicates the time of electrical stimulus, the vertical green and gray ticks indicate the time corresponding to the CCh or ACSF application during conditioning, correspondingly.

(D and E) SDFs of the neurons before (baseline) and after (post-cond) CCh conditioning (D) ( $n = 22$  neurons from 8 animals) or ACSF conditioning (E) ( $n = 20$  from 6 animals). For visualization, the color scale shows spike densities greater than 1 SD above spontaneous, normalized to the peak response magnitude during the baseline period. Cyan crosses represent neurons' calculated response durations (time to return to 1 STD of a neuron's spontaneous firing rate). Individual neuronal responses are sorted according to the baseline response durations.

(F) Population SDFs before (black) and after (green) CCh conditioning.

(G) Population SDFs before (black) and after (gray) control ACSF conditioning. The mean neuronal response duration before and after conditioning

is shown as the corresponding colored bars below the x axis. See also Figure S3.

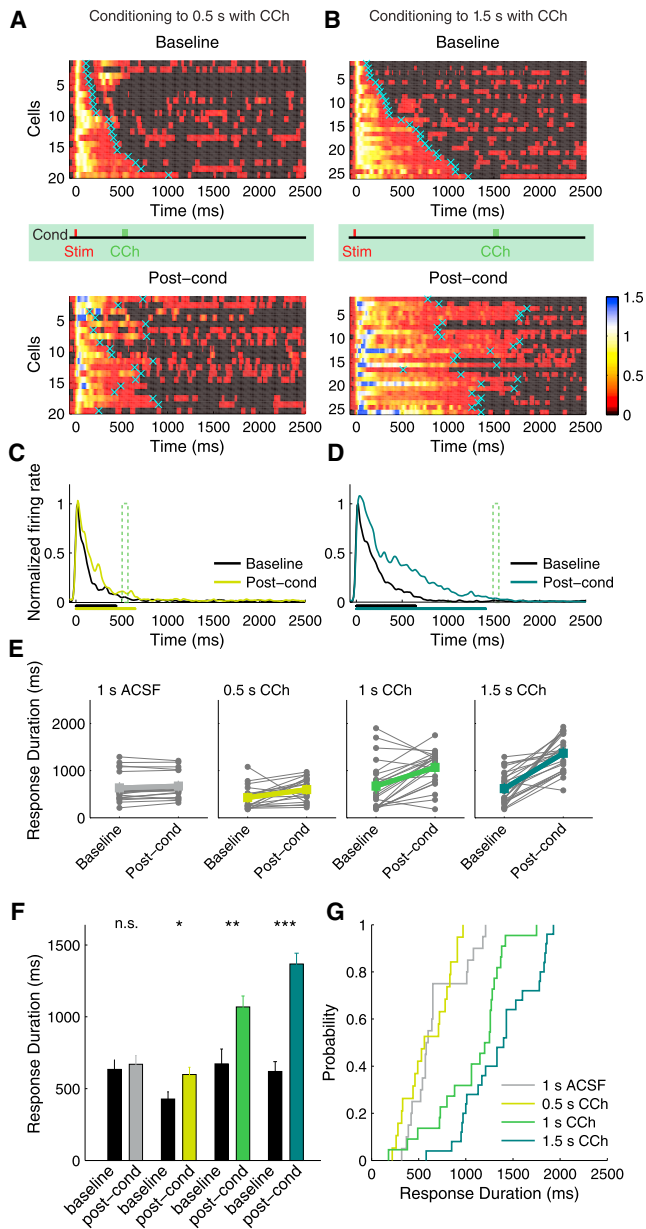
dominant cue went from being paired with a short delay to a long (Figure S1A, top) or a long delay to a short (Figure S1B, top). However, comparing the scores collected from 192-IgG-saporin-infused animals to the prereversal data demonstrates that 192-IgG-saporin-infused animals do not update their reports of time but rather continue to express the initial policy: cue 1 paired with a short delay (Figure S1A, bottom) and cue 2 paired with a long delay (Figure S1B, bottom).

While the NRTs obtained from 192-IgG-saporin-infused animals do not display a median shift after reversal, the shape of the distribution changes significantly (Kolmogorov-Smirnov test,  $p < 0.005$ ) due to an apparent increase in variance. This increased variability could either be a direct consequence of removing BF cholinergic innervation or occur in its prolonged absence. To differentiate between these possibilities, we categorized NRTs as belonging to an "early" (first three sessions after the reversal) or "late" (more than four sessions after) post-reversal epoch (Figure S2). In the early postreversal epoch, NRT update scores collected from intact and 192-IgG-saporin-infused animals are not significantly different from each other and are indistinguishable from the prereversal distribution. Whereas NRTs from intact animals eventually come to relate the new reward times given more experience with the new pairings, NRTs collected from 192-IgG-saporin-infused animals persist in reporting the outdated policy, albeit with diminished precision. Thus, the long-term consequences of cholinergic depletion after reversal are an inability to update temporal relationships and a degradation of the NRT population precision, possibly reflecting a neural signature of extinction.

### Teaching a Visual Cortical Slice to Report Time

Because the effects of 192-IgG-saporin treatment were restricted to the cholinergic innervation of the occipital cortex (Figures 1C–1E), the in vivo findings suggest that the plasticity underlying reward-timing activity occurs locally within V1. This insight encouraged us to attempt induction of response-duration plasticity (RDP) in acutely prepared slices of visual cortex, with electrical stimulation substituting for visual input and a brief pulse of the cholinergic agonist carbachol (CCh) serving as a proxy for the putative reinforcement signal. Because slice health and plasticity decline with age, we elected to use juvenile rather than adult animals for these experiments. Furthermore, we used mice rather than rats to establish an in vitro RDP protocol in anticipation of the need to use genetically engineered animals in future mechanistic investigations. This species difference does not undermine the relevance of the slice experiments, however, because mouse V1 also shows in vivo reward timing activity as observed in rat (C.-H. Liu and M.G.H.S., unpublished data).

We recorded extracellular spiking activity of layer 5 neurons in response to electrical stimulation (200  $\mu$ s) of the underlying white matter (WM) every 30 s (Figure 4A) in acutely prepared slices from male C57BL/6 mice (postnatal days 16–21). After collecting stable baseline responses, each slice was randomly assigned to a conditioning protocol involving either CCh or artificial cerebrospinal fluid (ACSF), which served as our control.



**Figure 5. Conditioning to Two Different Stimulus-CCh Delays, 0.5 s and 1.5 s**

(A) SDFs before and after conditioning to 0.5 s stimulus-reward interval;  $n = 19$  from 7 animals. Conventions are the same as Figure 4.  
 (B) Plots of SDFs before and after conditioning to 1.5 s stimulus-reward interval;  $n = 25$  from 10 animals. Individual neuronal responses are sorted according to the initial response time.  
 (C) Population SDFs conditioned to the 0.5 s stimulus-reward interval before (black) and after (light green) conditioning.  
 (D) Population SDFs conditioned to the 1.5 s stimulus-reward interval before (black) and after (dark green) conditioning. The mean neuronal response duration before and after conditioning is shown as the corresponding colored bars below the x axis.  
 (E) Connected filled circles represent response times of individual neurons before and after conditioning with ACSF (gray) and with CCh to different stimulus-reward times 0.5 s (light green), 1.0 s (green), and 1.5 s (dark green). Post-cond, postconditioning.  
 (F) Mean response times of the neurons before (black) and after (colored) conditioning. Error bars represent SEM. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ , Student's *t* test.  
 (G) Cumulative probability distributions of the neuronal response times after conditioning with ACSF and CCh. See also Figures S4 and S5.

Conditioning comprised local application of brief (30 ms) puffs of CCh (100  $\mu$ M) or ACSF to the preparation 1 s after each electrical stimulus for 200 trials. After conditioning, continuous perfusion with ACSF of slices was maintained for an additional 30 min. After this “washout,” we applied postconditioning WM stimuli to assess potential changes in the duration of electrically evoked spike trains.

Under our experimental conditions (see Experimental Procedures), each shock applied to the WM during baseline evoked a barrage of action potentials that usually subsided within a few hundred milliseconds (e.g., Figures 4B and 4C, top, and Figure S3). Although puffing ACSF 1 s after WM stimulation has no effect, CCh at that delay caused an increase after conditioning in the duration of the spike trains elicited by WM stimulation. We calculated spike density functions (SDFs) for each isolated single unit and averaged all SDFs within the same experimental condition before (baseline) and after conditioning to achieve population SDFs. For visualization, the SDFs of individual neurons before and after conditioning were normalized to the peak magnitude of their baseline population response (Figures 4D and 4E) and the population SDFs were compared (Figures 4F and 4G). To quantify the changes observed after conditioning on a per neuron basis, we credited the first moment of poststimulus time when the SDF was no longer above 1 STD of the neuron's spontaneous activity as being the neuron's response duration (cyan crosses in Figures 4D and 4E; see Experimental Procedures). A comparison of response durations before and after training reveals that the CCh-conditioned population has a substantially longer response (baseline =  $673 \pm 103$  ms, post-cond =  $1,069 \pm 76$  ms,  $p < 0.01$ , paired *t* test) terminating at a moment that accords well with the stimulus-CCh interval (Wilcoxon signed-rank test, no significant difference from 1 s stimulus-CCh interval,  $p = 0.13$ , but significant difference,  $p < 0.001$  and  $p < 0.0001$ , correspondingly, from 0.5 s and 1.5 s intervals that constitute alternative conditioning intervals tested below), an effect that was not observed after ACSF conditioning (baseline =  $634 \pm 67$  ms, post-cond =  $670 \pm 60$  ms,  $p > 0.05$ , paired *t* test; Wilcoxon signed-rank test,  $p < 0.001$  for 1 s).

While this first experiment showed that CCh conditioning can prolong WM-evoked responses, it did not assess whether the duration of conditioned responses is determined by the specific stimulus-CCh interval or is simply a nonspecific consequence of CCh pairing. To address this question, we performed a second experiment in which slices were conditioned using randomly assigned intervals of 0.5 s or 1.5 s (Figures 5A–5D). Response durations of the neurons conditioned with the 0.5 s stimulus-CCh interval increased significantly, though modestly, given the proximity of the mean baseline response to the target interval (baseline:  $428 \pm 50$  ms; post-cond:  $598 \pm 51$  ms,  $p < 0.01$ , paired *t* test). After conditioning with a stimulus-CCh interval of 1.5 s, the mean response duration increased from the baseline duration ( $619 \pm 69$  ms) to  $1,368 \pm 76$  ms ( $p < 0.001$ , paired *t* test). For

(F) Mean response times of the neurons before (black) and after (colored) conditioning. Error bars represent SEM. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ , Student's *t* test.

(G) Cumulative probability distributions of the neuronal response times after conditioning with ACSF and CCh. See also Figures S4 and S5.

comparison, response durations from all experimental groups before versus after conditioning are plotted in Figures 5E and 5F. Cumulative probability distributions allow direct comparison of response durations after conditioning with different stimulus-CCh intervals (Figure 5G). The distributions of response durations of neurons trained to the 1 s stimulus-CCh interval were significantly different from the response durations of neurons trained to both the 0.5 s interval and the 1.5 s interval (Kolmogorov-Smirnov test,  $p < 0.001$ ). Only after CCh conditioning were the central tendencies of the response durations significantly extended beyond their original baseline decay times to times that accorded well with their respective conditioned times (Wilcoxon signed-rank test, no significant difference from their corresponding stimulus-CCh intervals: 0.5 s –  $p = 0.10$ , 1.5 s –  $p = 0.23$ , but significant differences to the other intervals, 0.5 s: 1 s –  $p < 0.0001$ , 1.5 s –  $p < 0.0001$ ; 1.5 s: 0.5 s –  $p < 0.0001$ , 1 s –  $p < 0.001$ ).

These data show that timed CCh application can change the duration of evoked spike trains such that they relate to the stimulus-CCh interval. To determine whether CCh acts via muscarinic ACh receptors, we repeated the experiment with a 1 s stimulus-CCh interval in the presence of atropine (2  $\mu$ M). We found that this drug treatment completely blocked the induction of RDP (baseline:  $431 \pm 64$  ms; post-cond:  $388 \pm 56$  ms,  $p = 0.4267$ , Figure S4). Thus, muscarinic ACh receptor activation is necessary for the modifications responsible for RDP.

Stimulation of a visual cortical slice in the continuous presence of carbachol can lead to long-term potentiation (LTP) of evoked responses (Bröcher et al., 1992; Cho et al., 2012; but see Kirkwood et al., 1999). Although this mechanism alone seems unlikely to account for the precise timing of RDP after conditioning with different stimulus-CCh intervals, LTP of peak response magnitude offers one plausible mechanism for prolongation of an evoked spike train. Analysis of peak response magnitudes after conditioning did reveal an increase in the experiments using a 1.5 s delay (baseline:  $78 \pm 16$  spikes/s; post-cond:  $113 \pm 22$  spikes/s,  $p = 0.0014$ ), but not in experiments using the 0.5 s (baseline:  $71 \pm 12$  spikes/s; post-cond:  $82 \pm 13$  spikes/s) or 1 s (baseline:  $91 \pm 15$  spikes/s; post-cond:  $95 \pm 18$  spikes/s, Figure S5) delay intervals. To investigate this question further, we examined the relationship between changes in response duration and the corresponding changes in response magnitude on a cell-by-cell basis and found no correlation (Pearson's correlation coefficient  $R = -0.0606$ ,  $p = 0.5797$ ) (Figure S7A). We also examined the relationship between the change in response duration and the latency to the first spike and found no correlation (Pearson's correlation coefficient  $R = -0.0283$ ,  $p = 0.7958$ ) (Figure S7B). These results make LTP of monosynaptic inputs, directly activated by the electrical stimulation, an unlikely basis for RDP.

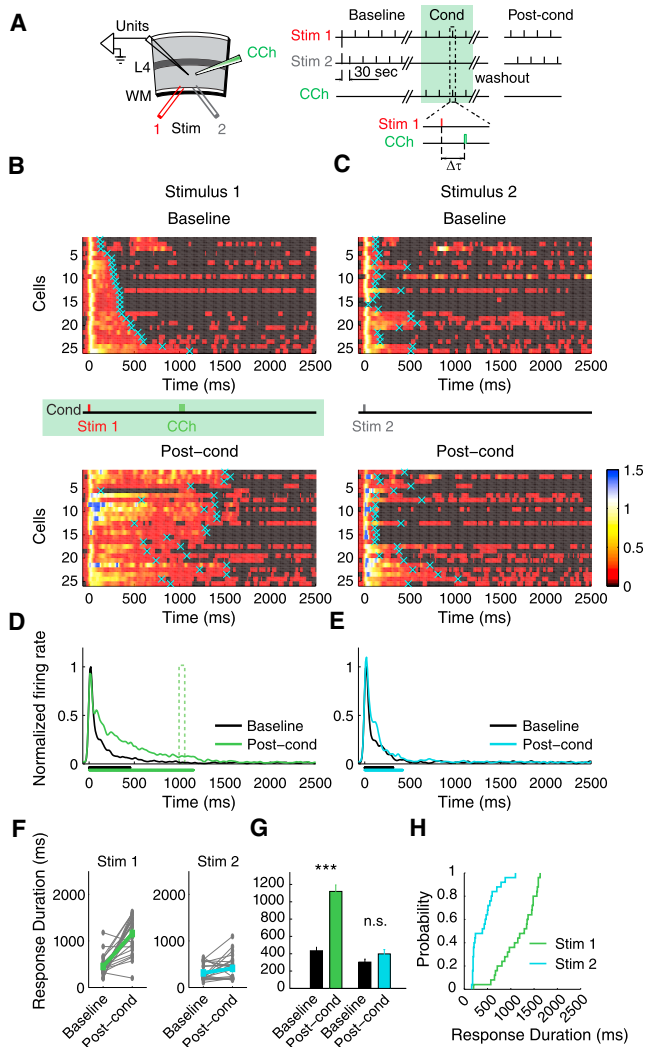
Another possibility is that CCh exerts its effects by strongly depolarizing neurons and evoking spikes at a fixed delay after WM stimulation. If this were the case, the coincidence of CCh-induced spiking with residual polysynaptic activity in layer 5 neurons could induce LTP of the active synapses and enhance reverberating activity. However, recording of spiking during conditioning revealed no such effect. Indeed, not only was CCh ineffective in evoking spikes, remarkably, the average

WM-evoked response durations did not change during conditioning (Figure S5). RDP only became apparent after washout of the CCh. To test whether a brief CCh puff produces sub-threshold depolarization, we conducted whole-cell patch-clamp recordings that revealed no direct effect on membrane potential at the time of puff (Figures S6A and S6B). To ensure that the training protocol was effective in these experiments, we continued conditioning the slices and confirmed, using extracellular recordings, successful entrainment to the 1 s stimulus-reward interval (Figure S6D).

One interesting property of reward timing activity in V1 in vivo is that individual neurons display cue dominance, reporting the stimulus-reward interval associated exclusively with one, but not both, visual cues (Shuler and Bear, 2006). To assess whether RDP in the slice also shows input specificity, we performed a two-pathway experiment in which only one pathway was conditioned (Figures 6A and 6B). We first collected baseline responses of neurons to two alternating stimulating electrodes (Stim 1 and 2) placed at a distance from each other in the white matter to allow for activation of at least partially nonoverlapping inputs that converge onto the recorded neurons. We then conditioned the slice by pairing only Stim 1 with a 100  $\mu$ M CCh puff applied at a 1 s delay (200 stimulus-CCh presentations). After 30 min washout in ACSF, we again collected responses from the same neurons to the two stimulating electrodes. Figure 6 shows, per neuron, the SDFs evoked by the conditioned (Stim 1, Figure 6B) and the nonconditioned (Stim 2, Figure 6C) electrodes before and after the pairing protocol, as well as the mean population response evoked by stimulating the conditioned (Figure 6D) or the nonconditioned (Figure 6E) pathway. Response durations pre- and postconditioning (Figure 6F) reveal that after conditioning, only the responses to the conditioned stimulus were extended to the time at which CCh was applied (baseline:  $450 \pm 44$  ms; post-cond:  $1,158 \pm 79$  ms,  $p < 0.001$ , paired t test), while the responses to the control stimulus were not significantly changed (baseline:  $314 \pm 35$  ms; post-cond:  $412 \pm 51$  ms,  $p = 0.055$ , paired t test) (Figures 6G and 6H). The cumulative distributions of the response durations evoked by the conditioned and the nonconditioned stimulating electrodes were also significantly different (Kolmogorov-Smirnov test,  $p < 0.001$ ) (Figure 6H).

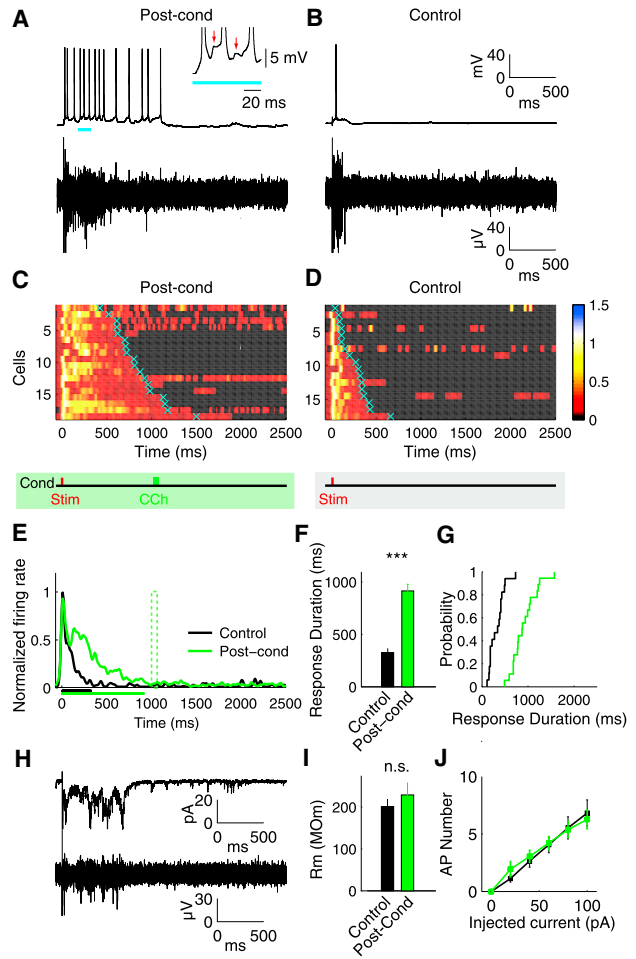
Theoretical and experimental studies suggest that alteration of neuronal timing, resulting from delayed behavioral reinforcement, can be accounted for either by synaptic modifications within a network of interconnected neurons (Gavornik and Shouval, 2011; Gavornik et al., 2009; Sanchez-Vives and McCormick, 2000) or by changes in the intrinsic electrical excitability of individual cells (Egorov et al., 2002; Fransén et al., 2006; Shouval and Gavornik, 2011). To decide which model better describes RDP in visual cortex, we performed whole-cell patch-clamp recordings of layer 5 neurons after conditioning slices with CCh at a 1 s delay between stimulus and “reward.” Neurons were patched in the vicinity of an extracellular recording electrode used to monitor baseline activity and the effect of conditioning. Current-clamp recordings of evoked action potentials confirmed RDP after CCh conditioning (Figures 7A–7D). Response durations of individual conditioned neurons were extended to the time when CCh had been applied after electrical





**Figure 6. Two-Pathway Experiment**

(A) Experimental setup. Stim 1 represents the conditioned stimulus; Stim 2 represents the control stimulus.  
 (B) SDFs before (baseline) and after (post-cond) conditioning with a 1 s stimulus-CCh interval,  $n = 25$  units from 9 animals. Conventions are the same as Figure 4.  
 (C) SDFs of the neurons in response to the control stimulus. Both (B) and (C) represent responses of the same neurons to Stim 1 and Stim 2; individual neuronal responses are sorted according to the initial duration of response to the Stim 1 electrode.  
 (D) Population SDF to the conditioned stimulus before (black) and after (green) conditioning.  
 (E) Population SDF to the control stimulus before (black) and after (cyan) conditioning. The mean neuronal response duration before and after conditioning is shown as the corresponding colored bars below the x axis.  
 (F) Individual response times of the conditioned (green) and the control (cyan) stimuli.  
 (G) Mean response times of the conditioned and the control stimuli,  $***p < 0.001$ ; n.s., not significant; Student's t test. Error bars represent SEM.  
 (H) Cumulative probability distributions of the neuronal responses after conditioning.



**Figure 7. Characterization of Response Time Plasticity Using Whole-Cell Patch Clamp**

(A) Representative traces of the current-clamp recordings (top) and the simultaneous multiunit recordings (bottom) postconditioning. The persistent firing is triggered by the barrages of synaptic potentials resulting in irregular spike frequency. The insert shows a part of the current-clamp recording with the arrows indicating excitatory postsynaptic potentials (EPSPs).  
 (B) Control traces.  
 (C and D) SDFs of the individual whole-cell recordings (C) postconditioning ( $n = 18$  from 10 animals) and from control whole-cell recordings (D) ( $n = 17$  from 8 animals). Neuronal responses are sorted according to their duration. Cyan crosses represent the calculated response durations.  
 (E) Population SDFs of the conditioned (green) and the control (black) neurons. The mean neuronal response duration for control and postconditioning is shown as the corresponding colored bars below the x axis.  
 (F) Mean response durations of the control (black) and the conditioned (green) neurons. Error bars represent SEM.  $***p < 0.001$ , Student's t test.  
 (G) Cumulative probability distributions of the control and the conditioned neuronal response durations.  
 (H) Voltage-clamp recording of the individual cell (top) and the simultaneous extracellular multiunit recording (bottom). Barrages of synaptic currents correspond well with the extracellular recordings of spiking activity.  
 (I) Membrane resistance of the control (black) and the conditioned (green) neurons was not significantly different. Error bars represent SEM.  
 (J) Intrinsic excitability of the conditioned and control neurons was also not significantly different as measured by the number of action potentials (APs) triggered by the increasing step current injections. Error bars represent SEM. See also Figures S6 and S7.

stimulation (conditioned:  $915 \pm 64$  ms, unconditioned control:  $325 \pm 40$  ms;  $p < 0.001$ , Student's *t* test) (Figures 7C–7F), and the distributions of response durations of the conditioned and control neurons were significantly different (Figure 7G; Kolmogorov-Smirnov test,  $p < 0.001$ ). Voltage-clamp experiments, in which the cell was held at the membrane potential  $V_m = -70$  mV, revealed a barrage of incoming synaptic currents corresponding well with extracellular recordings and the encoded stimulus-reward interval (Figure 7H). The mean synaptically evoked charge transfer was  $53.04 \pm 5.9$  pC ( $n = 8$ ) in conditioned neurons compared to  $15.53 \pm 2.7$  pC ( $n = 9$ ) in controls ( $p < 0.001$ , Student's *t* test). However, there was no difference in the membrane resistance of these same cells after conditioning (Figure 7I), nor in the function relating action potential number to the amount of current injected into the cell in current clamp (Figure 7J). Based on these findings, we conclude that changes in intrinsic membrane excitability of layer 5 neurons are unlikely to account for RDP.

## DISCUSSION

We find that selective removal of BF cholinergic input to V1 is sufficient to disrupt neural acquisition of cue-reward intervals while leaving intact the expression of intervals that were previously learned, indicating that V1 is the locus of plastic changes underlying reward timing activity and that ACh is necessary for this plasticity to occur. An instructive role for ACh in learning reward timing is indicated by the finding that pairing stimulation of V1 slice with a delayed pulse of the cholinergic agonist CCh extends the evoked spike train durations to relate the conditioned interval. Together, these data suggest that cholinergic input provides a signal of behavioral reinforcement to V1 and that even the reduced network contained within a visual cortical slice is sufficient to support reinforcement learning of temporal intervals as long as 1.5 s. The RDP protocol we describe here lays a foundation for future studies aimed at dissecting the mechanisms that underlie reward time prediction learning.

The neural process by which the brain attributes behavioral relevance to cues in the environment, and learns temporal expectancies based on that evidence, is unknown. Reinforcement learning theory provides a conceptual framework to explain how the brain might relate preceding neural activity with behavioral outcomes; by modifying relevant network elements on the basis of behavioral success, particular behaviors become more or less likely to occur in the future (Sutton and Barto, 1998). The biological implementation of this framework is achieved in part by invoking a global reinforcer, a signal that is permissive to plasticity and conveys the relative success or failure of recent actions throughout a network of neurons. However, relating the reinforcement signal directly to preceding neural activity presents a challenge for intervals that exceed the membrane time constant. Reinforcement learning theory contends with this “distal reward problem” by proposing the existence of an eligibility trace, a molecular memory of recent synaptic (or cellular) activity. Relating neural activity to future outcomes therefore requires an interaction between the reinforcement signal and slowly decaying eligibility traces, a notion that has been successfully applied to resolve the distal

reward problem (Izhikevich, 2007). However, the observed reward timing in V1 requires further explanation: not only are associations expressed relating cues to reward, but explicit temporal representations of expected cue-reward intervals are generated.

A formal model describing how visually cued expectancies can be learned and expressed within V1 has only recently been proposed (Gavornik et al., 2009). This model demonstrates how a reinforcement-based process termed “reward-dependent expression” generates not only the formation of appropriate associations, but also encodes the intervals interceding stimulus and reward. This model, and the *in vivo* results showing that ACh could serve as the reinforcement signal in V1, informed the design of our slice experiments. Consistent with theoretical assumptions, we found that when brief electrical stimulation of white matter is followed at a fixed delay by a pulse of CCh, the spike trains evoked in neurons by the electrical stimulation increase in duration to the time at which CCh was delivered. According to this theory, reward timing can be implemented either by adjusting the synaptic weights of a recurrent network of neurons (Gavornik et al., 2009) or by changing the intrinsic excitability of individual cells (Egorov et al., 2002; Fransén et al., 2006; Shouval and Gavornik, 2011). There are ample biological precedents for both types of change (Bekisz et al., 2010; Breton and Stuart, 2009; Budd, 1998; Johnson et al., 2010; Maravall et al., 2004; Nataraj et al., 2010; Rahman and Berger, 2011; Saar and Barkai, 2009). The RDP observed in layer 5 neurons appears to be better explained by the network model, as it is input specific and reflected in synaptic events without any changes in intrinsic membrane excitability. We cannot rule out changes in excitability in other neurons in the network, however.

The *in vitro* data presented here appear to correspond well with the results *in vivo*. We find that responses can be trained to represent time from stimulus to CCh “reward” with considerable accuracy and that the modified response is specific to the input that is associated with such reinforcement. However, there are some noteworthy differences. Perhaps the most apparent is that reward timing activity *in vivo* assumes one of three different forms: sustained increases in firing to the anticipated reward time, sustained decreases in firing to the anticipated reward time, and peak firing at the anticipated reward time (Figure 2). In slice, however, RDP manifested in neurons as sustained increases in the duration of spike trains evoked by WM stimulation. This difference is not accounted for by a species difference, as the same three classes of reward timing activity have been observed in mouse V1 (C.-H. Liu and M.G.H.S., unpublished data). It could be due to a different sample of neurons in the *in vivo* and *in vitro* experiments, since only layer 5 neurons were recorded in slices. Another possible explanation is that neurons in slices have much lower baseline firing rates than neurons *in vivo*, thus preventing detection of neurons with decreased firing to the time of anticipated CCh delivery, as well as detection of neurons that depend upon input from “sustained decrease” neurons to generate a response that peaks at the reward time. A final possibility is that RDP in the slice is a phenomenon that does not overlap mechanistically with the reward timing observed *in vivo*. However, given the critical involvement of cholinergic

transmission in both phenomena, we believe that the last explanation is unlikely.

The *in vivo* findings suggest that the release of ACh in V1 is a mechanistic requirement for acquisition of reward timing activity, and slice experiments show that CCh, acting via muscarinic ACh receptors, can elicit RDP to encode specific stimulus-CCh intervals. Precisely how CCh acts in this context as a reinforcement signal remains to be determined. One simple possibility for which there is precedence from cortical slice culture experiments (Johnson et al., 2010) is that CCh marks the time interval by triggering a second burst of action potentials, and by doing so triggers LTP of reverberating polysynaptic activity. However, our recordings during conditioning with CCh suggest that this is not the case. Under our experimental conditions, the CCh puff induces neither spiking nor subthreshold depolarization around the time of drug application. Our results are consistent with those of Gullledge et al. (2009). In that study, the authors applied brief pulses of a high concentration of ACh with pipettes in close proximity to layer 5 neurons in frontal cortex slices. The ACh did not evoke spikes directly and instead produced a phasic hyperpolarization with an  $\sim 1$  s delay. If the neurons were induced to fire action potentials with an intracellular current injection, the transient hyperpolarization was followed by an acceleration of the firing rate with a longer delay ( $\sim 2$  s). Both effects were blocked by atropine and were absent in the M1 receptor knockout. Because the neurons in our experiments were not firing action potentials at the time of the CCh application, there was no opportunity for spike acceleration. We did not observe CCh-induced hyperpolarization, possibly due to the fact that our application method was more diffuse and less focal. In any case, there is no support either from the literature or from our experiments that CCh initiates response duration plasticity by simply depolarizing layer 5 neurons at the time of application.

More work is required to identify the molecular basis of the theoretically postulated “eligibility traces” and to understand how conditioning results in encoding specific time intervals expressed as the duration of stimulus-elicited spike trains. Recently it has been shown that the sign of synaptic plasticity can change its polarity from LTP to LTD depending on the timing of cholinergic activation (Gu and Yakel, 2011; Pawlak et al., 2010). The visual cortical circuit might possess similar synaptic plasticity, and RDP may be its functional consequence at the network level. In any case, RDP in the slice offers an experimentally accessible preparation in which these behaviorally relevant questions can be addressed.

### Conclusion

Local disruption of cholinergic input is sufficient to block the acquisition of reward timing *in vivo*, and acutely isolated slices of V1 can learn to anticipate delivery of CCh. Collectively, these results indicate that this type of reward prediction can occur early in sensory processing streams and is not a specialized function allocated exclusively to higher-order cortical areas. In addition to introducing a robust *in vitro* protocol to study reinforcement learning, the present study advances the understanding of V1 reward timing activity and suggests that cholinergic neuromodulation serves as a reinforcement signal that informs the cortex about the outcome of behavior.

### EXPERIMENTAL PROCEDURES

All animal procedures followed NIH guidelines and were approved by the Johns Hopkins University IACUC or the Committee on Animal Care of Massachusetts Institute of Technology.

#### *In Vivo* Data Collection and Analysis

Sterile surgical procedures (see Supplemental Experimental Procedures) were used to chronically implant bilateral, movable electrode assemblies and infusion cannulae into binocular V1 (1–2 mm anterior, 4.5 mm lateral from lambda) of adult, male Long-Evans rats. Animals were housed in a vivarium with ad libitum food and, after recovery from surgery, given water only during behavioral sessions and brief scheduled access periods. Infusions were administered under light anesthesia and consisted of 350 nl saline or 192-IgG-saporin (13.5  $\mu\text{g/ml}$ ; Advanced Targeting Systems). After the final recording session, each animal was given a lethal dose of pentobarbital and transcardially perfused with cold PBS followed by 4% paraformaldehyde in PBS. After post-fixation, brain tissue was processed for acetylcholinesterase visualization (see Supplemental Experimental Procedures).

Before each session, animals were placed in an operant chamber containing a water spout within a nose poke port (Med Associates). Sixteen-channel headstage amplifiers were attached to electrode interface plugs overlying each hemisphere, and band-pass filtered (1–10 KHz) neural signals were sampled at 33 KHz by commercially available hardware (Neuralynx). Prior to recording, the spiking activity was inspected and a minute adjustment ( $\sim 25$   $\mu\text{m}$ ) of the electrode depth was applied if necessary to increase the yield or quality of recordings. Spike waveforms were sorted offline into single units (Offline Sorter, Plexon) by manually tracing boundaries between three-dimensional clusters formed by the following attributes: principal components 1 and 2, peak minus valley amplitude, nonlinear energy, and valley timing. Clusters that could not be cleanly separated from noise were excluded. The isolated timestamps were visualized with Neuroexplorer (Nex Technologies) and analyzed with MATLAB (MathWorks) scripts. Collapsing across all of the recording sessions from each animal (range: 12–40), each electrode yielded an average of  $3.031 \pm 0.887$  single-unit recordings.

Visual cues were presented through removable, head-mounted goggles overlying each eye and embedded with green LEDs. After a 2 s intertrial interval, the animal could initiate a trial by entering the nose poke, triggering a 100 ms flash to one of the goggles. The side of the cue presentation dictated the number of licks (range: 4–18) required to briefly activate a solenoid placed between the lick spout and a water reservoir. Rewarded and unrewarded trials were pseudorandomly interleaved such that no more than six consecutive trials could have the same stimulus and outcome. Behavioral sessions were controlled via custom signal conditioning hardware and MATLAB (MathWorks) software executed on the data acquisition computer.

For each single unit, cue dominance to the left eye (LE) or right eye (RE) cue was determined by comparing the correlations of the mean evoked responses in the poststimulus period ( $>100$  ms after stimulus offset up to average trial duration) to their differences. For instance, if the correlation between LE and LE-RE was greater than that of the correlation between RE and RE-LE, the neuron was classified as having cue dominance to the LE (and vice versa if the RE value was higher). Next, the area under the ROC curve was geometrically estimated bin-by-bin (10 ms, nonoverlapping) from the distributions formed by all trials of dominant (“signal”) and nondominant (“noise”) cue presentations (Wickens, 2001). Bias-corrected 95% confidence intervals were generated from 1,000 bootstrapped data sets drawn randomly from the signal and noise distributions with replacement (Obuchowski and Lieber, 1998). In the case of sustained inhibited neurons, the ROC area and confidence intervals were inverted. Cells were considered to have significantly different modulation if the lower bound of the confidence interval exceeded chance in more than 5% of the bins in the poststimulus period.

#### Brain Slice Experiments

Slices of visual cortex from male P16–P21 C57BL/6 mice were prepared as described in Philpot et al. (2001). Coronal slices (350  $\mu\text{m}$  thickness) recovered for 30 min at 32°C in a holding chamber filled with warmed

modified artificial cerebrospinal fluid (ACSF) that closely mimics ionic concentrations in physiological CSF (124 mM NaCl, 3.5 mM KCl, 1.25 mM Na<sub>2</sub>PO<sub>4</sub>, 26 mM NaHCO<sub>3</sub>, 0.8 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, and 10 mM dextrose, saturated with 95% O<sub>2</sub>, 5% CO<sub>2</sub>) (Di Terlizzi and Platt, 2006; Sachdev et al., 1998) and then continued to recover for an additional 2 hr at room temperature. All recordings were done in a submersion-type recording chamber maintained at 30°C and perfused at a rate of 2.5 ml/min with modified ACSF. Stimulation electrodes (clustered bipolar tungsten, FHC) were positioned in white matter and extracellular recordings were performed in layer 5 using glass recording electrodes ( $\approx 1$  M $\Omega$ ) filled with ACSF. Baseline and postconditioning responses were collected every 30 s, and the stimulus intensity (30–150  $\mu$ A, 0.2 ms duration) was chosen to achieve a stable persistent response duration longer than 30 ms. For conditioning, we used a Picospritzer II (General Valve) to apply brief pulses of CCh to a slice (30 ms puffs of 100  $\mu$ M CCh).

Extracellular recordings made using Axopatch 200B (Axon Instruments) were amplified 1,000 times, filtered between 0.3 and 3 kHz, and digitized at 10 or 30 kHz. Offline spike detection and discrimination of single-unit activity was done using stochastic wavelet analysis, with spike events detected using amplitude thresholding at 3.5 $\times$  or 5 $\times$  the SD of the noise (MATLAB [MathWorks]; software by Rodrigo Quiroga, <http://www.vis.caltech.edu/~rodri>) (Quiroga et al., 2004). Spike density functions (SDFs) were built from the peristimulus time histograms with 10 ms bins and smoothed by convolution with a Gaussian kernel ( $\sigma = 70$  ms). The last ten SDFs of baseline and the first ten SDFs postconditioning (after the “washout” period) were averaged for each neuron. The effect of conditioning was visualized at the population, as well as the individual neuron level. At the population level, the responses of all neurons in each experimental condition were averaged to create population SDFs before and after training

$$sdf_{population} = \frac{\sum_{i=1}^N sdf_i}{N}$$

Population SDFs were then normalized to the peak magnitude of the population SDF before training (baseline). SDFs of individual neurons were similarly normalized for visualization. To calculate each individual neuron's response duration, we first subtracted the neuron's spontaneous activity +1 SD from its SDF and then defined the neuron's response duration as the time of the first zero crossing of the resulting SDF. These neuronal response durations were then compared before and after conditioning.

Whole-cell patch-clamp recordings were done using pipettes with open tip resistances 3–5 M $\Omega$ . For current-clamp recordings, the pipette internal solution contained 20 mM KCl, 100 mM Na-gluconate, 10 mM HEPES, 4 mM MgATP, 0.3 mM Na<sub>2</sub>GTP, 7 mM phosphocreatine-Tris, and 0.2% biocytin with pH adjusted to 7.2 and osmolarity adjusted to 300 mOsm; for voltage clamp, the internal solution contained 103 mM Cs-gluconate, 5 mM TEA-Cl, 2.8 mM NaCl, 20 mM HEPES, 0.3 mM Na<sub>2</sub>GTP, 4 mM MgATP, 10 mM Na<sub>2</sub>-phosphocreatine, 0.2 mM EGTA, 0.2% biocytin, and 5 mM QX-314-Cl with pH adjusted to 7.2 and osmolarity adjusted to 300 mOsm with sucrose or ddH<sub>2</sub>O. Seal resistances were >1 G $\Omega$ , and pipette capacitive transients were minimized prior to breakthrough. Only cells with series resistance <30 M $\Omega$  were included in this study. Membrane resistance and cell excitability were measured in response to a series of incremental 200 ms step current injections from –100 pA to 100 pA. Whole-cell recordings were made using Axopatch 200B (Axon Instruments). Evoked synaptic charge transfer was calculated as the integral of the postsynaptic currents over time from the moment after electrical stimulation to 1.5 s poststimulus.

To prevent experimental bias, we performed all experiments in pairs of protocol conditions in which the type of experiment was determined by a coin flip. The data are reported as means  $\pm$  SEM. To determine the statistical significance of the difference in response durations before and after conditioning, we used paired Student's *t* test. To compare the difference in response durations posttraining for different conditioning intervals and for the two-pathway experiments, we used the Kolmogorov-Smirnov test. All *in vitro* analyses were done using pClamp (Axon Instruments) and custom-written MATLAB and python scripts.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.neuron.2012.12.039>.

## ACKNOWLEDGMENTS

We thank Harel Shouval, Jeff Gavornik, and Arnold Heynen for insightful comments and critiques. We are grateful to Jill McGaughy, Weidong Hu, and Peter Holland for sharing their histological expertise. We thank Tanya Marton, Vijay M.K. Nambodiri, and Camila Zold for helpful discussions of the experimental design. We also thank Hannah Joo and Shelby Hecht for their assistance with data collection, electrode drive assembly, and histology and Kathleen Oram, Erik Sklar, and Suzanne Meagher for technical and administrative support. This work was funded by NIMH (R01 MH084911) and start-up funds provided by the Johns Hopkins University to M.G.H.S., NIDA (F31 DA026687) to E.B.R., and grants from NEI (R01 EYO12309) and NICHD (R01 HD046943) to M.F.B.

Accepted: December 21, 2012

Published: February 20, 2013

## REFERENCES

- Bear, M.F., and Singer, W. (1986). Modulation of visual cortical plasticity by acetylcholine and noradrenaline. *Nature* 320, 172–176.
- Bekisz, M., Garkun, Y., Wabno, J., Hess, G., Wrobel, A., and Kossut, M. (2010). Increased excitability of cortical neurons induced by associative learning: an *ex vivo* study. *Eur. J. Neurosci.* 32, 1715–1725.
- Breton, J.-D., and Stuart, G.J. (2009). Loss of sensory input increases the intrinsic excitability of layer 5 pyramidal neurons in rat barrel cortex. *J. Physiol.* 587, 5107–5119.
- Bröcher, S., Artola, A., and Singer, W. (1992). Agonists of cholinergic and noradrenergic receptors facilitate synergistically the induction of long-term potentiation in slices of rat visual cortex. *Brain Res.* 573, 27–36.
- Budd, J.M. (1998). Extrastriate feedback to primary visual cortex in primates: a quantitative analysis of connectivity. *Proc. Biol. Sci.* 265, 1037–1044.
- Carey, R.G., and Rieck, R.W. (1987). Topographic projections to the visual cortex from the basal forebrain in the rat. *Brain Res.* 424, 205–215.
- Cho, K.H., Jang, H.J., Jo, Y.H., Singer, W., and Rhie, D.J. (2012). Cholinergic induction of input-specific late-phase LTP via localized Ca<sup>2+</sup> release in the visual cortex. *J. Neurosci.* 32, 4520–4530.
- Conner, J.M., Culberson, A., Packowski, C., Chiba, A.A., and Tuszynski, M.H. (2003). Lesions of the Basal forebrain cholinergic system impair task acquisition and abolish cortical plasticity associated with motor skill learning. *Neuron* 38, 819–829.
- Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Curr. Opin. Neurobiol.* 18, 185–196.
- Di Terlizzi, R., and Platt, S. (2006). The function, composition and analysis of cerebrospinal fluid in companion animals: part I - function and composition. *Vet. J.* 172, 422–431.
- Doya, K. (2002). Metalearning and neuromodulation. *Neural Netw.* 15, 495–506.
- Egorov, A.V., Hamam, B.N., Fransén, E., Hasselmo, M.E., and Alonso, A.A. (2002). Graded persistent activity in entorhinal cortex neurons. *Nature* 420, 173–178.
- Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Fransén, E., Tahvildari, B., Egorov, A.V., Hasselmo, M.E., and Alonso, A.A. (2006). Mechanism of graded persistent cellular activity of entorhinal cortex layer V neurons. *Neuron* 49, 735–746.
- Froemke, R.C., Merzenich, M.M., and Schreiner, C.E. (2007). A synaptic memory trace for cortical receptive field plasticity. *Nature* 450, 425–429.

- Gavornik, J.P., and Shouval, H.Z. (2011). A network of spiking neurons that can represent interval timing: mean field analysis. *J. Comput. Neurosci.* *30*, 501–513.
- Gavornik, J.P., Shuler, M.G., Loewenstein, Y., Bear, M.F., and Shouval, H.Z. (2009). Learning reward timing in cortex through reward dependent expression of synaptic plasticity. *Proc. Natl. Acad. Sci. USA* *106*, 6826–6831.
- Genovesio, A., Tsujimoto, S., and Wise, S.P. (2009). Feature- and order-based timing representations in the frontal cortex. *Neuron* *63*, 254–266.
- Gu, Q. (2003). Contribution of acetylcholine to visual cortex plasticity. *Neurobiol. Learn. Mem.* *80*, 291–301.
- Gu, Q., and Singer, W. (1993). Effects of intracortical infusion of anticholinergic drugs on neuronal plasticity in kitten striate cortex. *Eur. J. Neurosci.* *5*, 475–485.
- Gu, Z., and Yakes, J.L. (2011). Timing-dependent septal cholinergic induction of dynamic hippocampal synaptic plasticity. *Neuron* *71*, 155–165.
- Gulledge, A.T., Bucci, D.J., Zhang, S.S., Matsui, M., and Yeh, H.H. (2009). M1 receptors mediate cholinergic modulation of excitability in neocortical pyramidal neurons. *J. Neurosci.* *29*, 9888–9902.
- Holley, L.A., Wiley, R.G., Lappi, D.A., and Sarter, M. (1994). Cortical cholinergic deafferentation following the intracortical infusion of 192 IgG-saporin: a quantitative histochemical study. *Brain Res.* *663*, 277–286.
- Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* *160*, 106–154.
- Izhikevich, E.M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* *17*, 2443–2452.
- Johnson, H.A., Goel, A., and Buonomano, D.V. (2010). Neural dynamics of in vitro cortical networks reflects experienced temporal patterns. *Nat. Neurosci.* *13*, 917–919.
- Kilgard, M.P., and Merzenich, M.M. (1998). Cortical map reorganization enabled by nucleus basalis activity. *Science* *279*, 1714–1718.
- Kirkwood, A., Rozas, C., Kirkwood, J., Perez, F., and Bear, M.F. (1999). Modulation of long-term synaptic depression in visual cortex by acetylcholine and norepinephrine. *J. Neurosci.* *19*, 1599–1609.
- Leon, M.I., and Shadlen, M.N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* *38*, 317–327.
- Lin, S.-C., and Nicolelis, M.A. (2008). Neuronal ensemble bursting in the basal forebrain encodes salience irrespective of valence. *Neuron* *59*, 138–149.
- Lysakowski, A., Wainer, B.H., Bruce, G., and Hersh, L.B. (1989). An atlas of the regional and laminar distribution of choline acetyltransferase immunoreactivity in rat cerebral cortex. *Neuroscience* *28*, 291–336.
- Maravall, M., Stern, E.A., and Svoboda, K. (2004). Development of intrinsic properties and excitability of layer 2/3 pyramidal neurons during a critical period for sensory maps in rat barrel cortex. *J. Neurophysiol.* *92*, 144–156.
- Mechawar, N., Cozzari, C., and Descarries, L. (2000). Cholinergic innervation in adult rat cerebral cortex: a quantitative immunocytochemical description. *J. Comp. Neurol.* *428*, 305–318.
- Mita, A., Mushiaki, H., Shima, K., Matsuzaka, Y., and Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nat. Neurosci.* *12*, 502–507.
- Myers, R.D. (1966). Injection of solutions into cerebral tissue: Relation between volume and diffusion. *Physiol. Behav.* *1*, 171–174, IN9.
- Nataraj, K., Le Roux, N., Nahmani, M., Lefort, S., and Turrigiano, G. (2010). Visual deprivation suppresses L5 pyramidal neuron excitability by preventing the induction of intrinsic plasticity. *Neuron* *68*, 750–762.
- Obuchowski, N.A., and Lieber, M.L. (1998). Confidence intervals for the receiver operating characteristic area in studies with small samples. *Acad. Radiol.* *5*, 561–571.
- Origlia, N., Kuczewski, N., Pesavento, E., Aztiria, E., and Domenici, L. (2008). The role of cholinergic system in neuronal plasticity: focus on visual cortex and muscarinic receptors. *Arch. Ital. Biol.* *146*, 165–188.
- Pawlak, V., Wickens, J.R., Kirkwood, A., and Kerr, J.N. (2010). Timing is not everything: neuromodulation opens the STDP gate. *Front Synaptic Neurosci.* *2*, 146.
- Paxinos, G., and Watson, C. (2008). *The Rat Brain in Stereotaxic Coordinates* (London: Academic Press).
- Pennartz, C.M. (1995). The ascending neuromodulatory systems in learning by reinforcement: comparing computational conjectures with experimental findings. *Brain Res. Brain Res. Rev.* *21*, 219–245.
- Philpot, B.D., Sekhar, A.K., Shouval, H.Z., and Bear, M.F. (2001). Visual experience and deprivation bidirectionally modify the composition and function of NMDA receptors in visual cortex. *Neuron* *29*, 157–169.
- Quiroga, R.Q., Nadasdy, Z., and Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* *16*, 1661–1687.
- Rahman, J., and Berger, T. (2011). Persistent activity in layer 5 pyramidal neurons following cholinergic activation of mouse primary cortices. *Eur. J. Neurosci.* *34*, 22–30.
- Rye, D.B., Wainer, B.H., Mesulam, M.M., Mufson, E.J., and Saper, C.B. (1984). Cortical projections arising from the basal forebrain: a study of cholinergic and noncholinergic components employing combined retrograde tracing and immunohistochemical localization of choline acetyltransferase. *Neuroscience* *13*, 627–643.
- Saar, D., and Barkai, E. (2009). Long-lasting maintenance of learning-induced enhanced neuronal excitability: mechanisms and functional significance. *Mol. Neurobiol.* *39*, 171–177.
- Sachdev, R.N.Â.S., Lu, S.-M., Wiley, R.G., and Ebner, F.F. (1998). Role of the basal forebrain cholinergic projection in somatosensory cortical plasticity. *J. Neurophysiol.* *79*, 3216–3228.
- Sanchez-Vives, M.V., and McCormick, D.A. (2000). Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat. Neurosci.* *3*, 1027–1034.
- Santos-Benitez, H., Magariños-Ascone, C.M., and Garcia-Austt, E. (1995). Nucleus basalis of Meynert cell responses in awake monkeys. *Brain Res. Bull.* *37*, 507–511.
- Shouval, H.Z., and Gavornik, J.P. (2011). A single spiking neuron that can represent interval timing: analysis, plasticity and multi-stability. *J. Comput. Neurosci.* *30*, 489–499.
- Shuler, M.G., and Bear, M.F. (2006). Reward timing in the primary visual cortex. *Science* *311*, 1606–1609.
- Sutton, R.S., and Barto, A.G. (1998). *Reinforcement Learning* (Cambridge, MA: MIT Press).
- Tago, H., Kimura, H., and Maeda, T. (1986). Visualization of detailed acetylcholinesterase fiber and neuron staining in rat brain by a sensitive histochemical procedure. *J. Histochem. Cytochem.* *34*, 1431–1438.
- Waite, J.J., Wardlow, M.L., Chen, A.C., Lappi, D.A., Wiley, R.G., and Thal, L.J. (1994). Time course of cholinergic and monoaminergic changes in rat brain after immunolesioning with 192 IgG-saporin. *Neurosci. Lett.* *169*, 154–158.
- Weinberger, N.M. (2003). The nucleus basalis and memory codes: auditory cortical plasticity and the induction of specific, associative behavioral memory. *Neurobiol. Learn. Mem.* *80*, 268–284.
- Wickens, T.D. (2001). *Elementary Signal Detection Theory* (Oxford: Oxford University Press).
- Wiley, R.G., Oeltmann, T.N., and Lappi, D.A. (1991). Immunolesioning: selective destruction of neurons using immunotoxin to rat NGF receptor. *Brain Res.* *562*, 149–153.
- Wilson, F.A., and Rolls, E.T. (1990). Neuronal responses related to reinforcement in the primate basal forebrain. *Brain Res.* *509*, 213–231.
- Woolf, N.J. (1996). Global and serial neurons form a hierarchically arranged interface proposed to underlie memory and cognition. *Neuroscience* *74*, 625–651.
- Zilles, K. (1985). *The Cortex of the Rat: A Stereotaxic Atlas* (Berlin: Springer-Verlag).