SUPPLEMENTARY METHODS

Animals

All experiments were performed in compliance with the regulations of the Animal Care and Use Committees at the University of California, Berkeley and at the NIAAA, and according to NIH guidelines. Six male Long-Evans rats weighing roughly 250 grams were used for the experiments. In addition, seven striatal-specific NMDAR1-knockout mice and eight littermate controls generated by crossing *RGS9L-cre* mice with *NMDAR1-loxP* mice as described previously[1] were used in the BMI experiments. Behavioral experiments for supplemental figure 11 were performed on an additional group of *RGS9L-cre+/NMDAR1-loxP* homozygous mice and controls consisting of either *RGS9L-cre+ carrying no NMDAR1-loxP floxed alleles* or *RGS9L-cre-/NMDAR1-loxP* homozygous mice.

Surgery

Rodents were chronically implanted with microwire arrays ipsilaterally in M1 and the DS. In rats 2 arrays were independently implanted: each array contained 16 tungsten microelectrodes (35 μm diameter, 250 μm electrode spacing, 8x2 configuration; Innovative Neurophysiology, Durham, NC). Stereotactic coordinates relative to bregma were used to center the arrays. In rats, these coordinates were: anteroposterior 2 mm, mediolateral 2 mm, and dorsoventral 1.5 mm for M1; and anteroposterior 0.5 mm, mediolateral 4 mm, dorsoventral 4.5 mm for DLS. In mice, a customized array contained 32 tungsten microeletrodes (8 columns x 4 rows configuration, 35 μm diameter, 150 μm electrode spacing, 200-1000-200 μm electrode row spacing, 1.5 mm shorter for medial than lateral two rows; Innovative Neurophysiology, Durham,

NC) was implanted unilaterally, with the medial two rows of electrodes targeting M1 and the lateral two rows of electrodes targeting DS; coordinates centered at anteroposterior +0.5 mm and mediolateral +2.0 mm, lowering dorsoventrally from brain surface 1.0-1.1 mm for M1 and 2.2 mm for DS). In all cases, M1 implants were targeted to record from layer 5 pyramidal neurons. Rats were anesthetized with Ketamine (50 mg/kg) and Xylazine (5 mg/kg) with supplemental isoflurane gas as needed. Mice were anesthetized with isoflurane. Craniotomies were sealed with cyanoacrylate and rodents were allowed to recover for ten days after implantation before behavioral training. Rats were given dexamethasone treatment (0.5 mg/kg) for one week following surgery to minimize tissue damage around the implant[2].

Two male Long-Evans rats were also chronically implanted with subdermal Teflon-coated Tungsten microwires (Grass Technologies, West Warwick, RI) to record the differential EMG of the intrinsic muscles of the mystacial pad. An incision was made in the skin at the midline of the skull. A 25-gauge syringe needle was loaded with an electrode and inserted into the muscular tissue of the mystacial pad, and four individual electrodes were spaced evenly along the extent of the intrinsic musculature. Placement of the electrodes was verified by passing trains of 200 μA monophasic pulses lasting 200 μs with an interval of 1 ms through the electrodes and observing evoked movements of the vibrissae.

Electrophysiology

Single unit activity and local field potentials were simultaneously recorded with a Multichannel Acquisition Processor (MAP; Plexon Inc., Dallas, TX). Activity was sorted using an online sorting application (Plexon Inc., Dallas, TX) prior to each daily recording session. Only units with a clearly identified waveform and high signal-to-noise ratio were used. Sorting

templates were further refined using an offline sorting application (Plexon Inc., Dallas, TX). Behavioral timestamps were sent to the MAP recording system through Matlab (Mathworks, Natick, MA) and synchronized to the neural data for later analyses. The differential EMG was recorded simultaneously using a DAM50 Differential Amplifier (World Precision Instruments, Inc., Sarasota, FL).

Behavioral Task

After rodents recovered from surgery, two ensembles of 2-4 well-isolated M1 units each were chosen for testing based on waveforms, interspike-interval histograms, and refractory periods. Activity from these ensembles was binned in 200 millisecond bins and entered into an online transform algorithm that related neural activity to the pitch of an auditory cursor, and by modulating activity in these ensembles, rodents controlled the pitch of the cursor. The specific transform used was:

$$f(t) = A_1 e^{(\Sigma_{i \in N_1} r_i(t))} - A_2 e^{(\Sigma_{j \in N_2} r_j(t))} + B$$

where f is the cursor frequency, $r_i(t)$ is the firing rate for neuron i at time bin t, N_1 and N_2 denote the units in ensembles one and two, respectively, and A_1 , A_2, and B are coefficients that are set based on a daily baseline recording session of 250 cycles. To impose smoothness on the cursor position, a moving average of 3 time bins was used. This transform caused increased activity in the first ensemble to produce increases in the cursor pitch, while increased activity in the other ensemble produced decreases in the cursor pitch. Linear changes in firing rate resulted in exponential changes in cursor frequency, and frequency changes were binned in quarter-octave intervals to match rodent psychophysical discrimination thresholds[3]. The chosen ensembles and transform coefficients were updated daily in accordance with signal quality.

However, the same or nearby M1 units were maintained for use in BMI control whenever the signal quality permitted.

The rodents had to then precisely modulate these neuronal ensembles to move the cursor to one of two target pitches, one which was associated with a 20% (10% in mice) sucrose solution reward and one which was associated with a 45 mg (20 mg in mice) food pellet reward. Rodents were free to choose either reward at any time, although M1 activity levels had to return to baseline levels for a new trial to begin. A trial was marked incorrect if neither of these target states were achieved within 30 seconds of trial initiation. Chance levels of target achievement were assessed by collecting daily baseline recordings in which neural activity was entered into the transform algorithm but rodents were not engaged in the task and received no auditory feedback to guide performance.

To control for physical movements, rats also performed the task after being given injections of lidocaine in the whisker pad to locally inactivate sensory and motor nerve endings[4]. Array placement in the vibrissa region of M1 was confirmed by applying trains of 60 μA biphasic pulses through the recording array and observing subsequent twitches of the whisker pad. Pulses lasted 200 μs and were delivered at 350 Hz. 0.5-1 mL of a 2 mg/mL lidocaine solution were then injected into the whisker pad immediately before the behavioral session.

Behavioral sessions took place in a rodent operant box equipped with liquid and pellet dispensers (For rats: Lafayette Instrument Company, Lafayette, IL; For mice: Med Associates, Inc., St. Albans, VT). Recorded neural data was entered in real time to custom routines in Matlab that then translated those activity levels into the appropriate feedback pitch and played the pitch on speakers mounted on 2 sides of the operant box. Frequencies used for auditory feedback ranged from 1-24 kHz in quarter-octave increments. When a target was hit, a Data

Acquisition board (National Instruments, Austin, TX) controlled by Matlab triggered the operant box to supply the appropriate reward to rodents.

For lever-pressing operant training, *RGS9L-Cre/Nr1*$^{f/f}$ animals and their littermate controls were firstly trained to acquire a regular sequence task under fixed-ratio eight (FR8) schedule for one week as described before[5]. After that, a differential reinforcement sequence training schedule was introduced where only eight lever presses performed within 16 seconds will be reinforced. Self-paced lever press sequences were defined by either the statistics of lever pressing for each animal (bimodal or Poisson distribution) or by a bout of licks interrupting lever pressing. The sequence length and duration were thus calculated based on each individual sequence, and the within-sequence press rate was computed by the ratio of sequence length (≥ 2 presses) and the corresponding sequence duration[5].

Action-Outcome Manipulations

After initial training, the action-outcome contingency was degraded for 2 days. M1 ensemble activity still determined the pitch of the cursor, but reward was now given on a variable time schedule with equal probability of getting a reward after target achievement and no target achievement. Contingency was then reinstated for 2 days; performance had to return to previous levels before further task manipulations were performed.

A sensory-specific satiety test was also performed to manipulate expected reward value – Devaluation test. Animals were given free access to one of the task rewards in a cage different than their home cage for one hour before a choice session. Animals were then placed into the operant training box and allowed to perform the task with free choice over which target to hit at any time during the session.

Finally, an omission test was performed to demonstrate that rodents were able to intentionally inhibit the learned neuroprosthetic actions. For this test, animals were no longer rewarded after achieving one of the behavioral targets and the reward associated with that target was now delivered when animals successfully inhibited the action for the duration of the trial (30s). During this manipulation, the other target continued to be rewarded normally. This manipulation was performed separately for both targets.

Data Analysis

All analyses were performed in Matlab (Mathworks, Natick, MA) with custom-written programs. Unit data were first binned in 1 millisecond time bins and digitized. As no difference was seen between the two targets in the basic task, data from the two targets were pooled. Peri-stimulus time histograms (PSTHs) were calculated in relation to target achievement and firing rate analyses were performed from 4 seconds before until 4 seconds after this event, as well as from 6 seconds before until 1 second after this event. Firing rates and PSTHs were smoothed with a moving window of 100 milliseconds. For target-related firing rate modulation, similar to the data analysis previously applied to relate natural movements to neural activity[5-7], distributions of PSTHs from -5 to -2 s (before target reaching) were considered baseline activity. A significant increase in firing rate was defined if at least 20 consecutive bins had firing rates larger than a threshold of 99% above baseline activity, and a significant decrease in firing rate was defined if at least 20 consecutive bins had a firing rate smaller than a threshold of 95% below baseline activity[5-7]. Representative days for early and late learning were chosen for each animal from days 2-4 and 8-11, respectively, based on behavioral performance, motivation, and signal quality. For firing rate z-score analyses, firing rates for individual units were binned in

100 millisecond time bins, averaged across events, and smoothed with a moving average. Mean firing rates for each unit were then z-scored and plotted during both early and late learning sessions. Cross-correlograms were also calculated with spiking activity in either M1 or DS being correlated to the occurrence of action potentials in the other region.

For coherence analyses, a multi-taper method was used to compute spectral estimates of spiking activity in both regions[8-10]. A total of 5 tapers were used and estimates were computed every 50 milliseconds with a window size of 500 milliseconds. Coherence between spiking activity in the two regions was calculated and defined defined as:

$$C_{xy} = \frac{|R_{xy}|}{\sqrt{|R_{xy}|}\sqrt{|R_{yy}|}}$$

where R_{xx} and R_{yy} are the power spectra and R_{xy} is the cross-spectrum. Power and spike-spike coherence estimates were calculated relative to the delivery of reward and averaged across trials and animals. Mean coherence in the theta band was calculated and defined as 4-8 Hz (4.5-9 Hz in mice). It has been shown that values for spike-spike coherence are generally lower than values for spike-field or field-field coherence, and these values are on scale with those seen in the current work[11] Because there were a different number of correct trials in early and late learning, coherence in late learning was calculated based on a randomly-selected subset of trials matched to the number of trials in early learning.

EMG data were analyzed by first high-pass filtering the raw time series at 10 Hz and then calculating the absolute value of the Hilbert transform of this filtered signal[12]. Correlations between EMG and spiking activity were calculated after smoothing both signals with a moving average of 50 msec.

The signal-to-noise ratio for each recorded waveform was quantified as:

$$SNR = \frac{A}{2 * SD_{noise}}$$

where A is the peak-to-peak voltage of the mean waveform and SD_{noise} is the standard deviation of the residuals from each waveform after the mean waveform has been subtracted[13].

The burst-like phasic activity of striatal MSNs was defined as two or more spikes occurring with an inter-spike interval of less than 125 ms and terminated with an inter-spike interval more than 280 ms[5].

Histology

Final placement of the electrodes was monitored online during surgery based on neural activity, and then confirmed histologically at the end of experiments. After the completion of experiments, rodents were deeply anesthetized with sodium pentobarbital (150 mg/kg) and electrolytic lesions were made at electrode tips by passing current (10 μA, 15 sec) through the recording electrodes. Rats were transcardially perfused with saline followed by a 15% formaldehyde solution. Brains were removed and fixed in a solution of 10% sucrose and 15% formaldehyde. 100 μm coronal sections were then made through the forebrain with a vibratome. Sections were mounted on gelatin-coated slides and stained with cresyl violet. In mice, brains were perfused with 10 % formalin, fixed in a solution of 30 % sucrose and 10 % formalin, sectioned with a cryostat (coronal slices of 40 - 60 μm), and stained either with cresyl violet or with DAPI (4',6-diamidino-2-phenylindole, 0.25 mg/L into PBS) and Propidium iodide (0.2 mg/L into PBS).

Immunohistochemistry

RGS9L-cre mice were crossed with Gt(ROSA)26Sortm1(Smo/EYFP)Amc/J (ROSA-YFP; Jackson Laboratory) to verify in which cells Cre deletes in this line. Adult mice were anaesthetized (IsoFluorane) and transcardially perfused with 0.9% saline followed by 4% buffered paraformaldehyde. Brains were cut at 50 μm on a Vibratome (Leica) and sections were subsequently rinsed in PBS (Phosphate-buffered saline, pH 7.2) several times prior to overnight incubation in primary antibodies diluted in PBS-0.1% Triton X-100 (v/v; Sigma, St. Louis, MO). Antibodies used included mouse anti-parvalbumin (1:10000, Sigma) and Alexa488 anti-GFP (1:1000, Molecular Probes) for YFP enhancement. The next day, sections were washed five times in PBS and then incubated for 2 h in Alexa 594-anti-mouse secondary antibody (Molecular Probes) diluted (1:3000) in PBS – 0.1% Triton X-100. In the last wash, sections were counterstained with 4,6-diamidino-2-phenylindole (DAPI 1:1000, Sigma) to visualize cell nuclei and washed sections were placed on Superfrost glass slides (Fisher) and coversliped with Mowiol (Sigma). Images were obtained using AxioVision MosaiX software on a motorized Zeiss Imager.M2 microscope with an AxioCam MRm (Zeiss) camera. Images were saved as tiff files and processed in Photoshop CS4.

Pharmacology

NMDA antagonist MK-801 (0.1-0.3 mg/kg; Dizocilpine, Sigma-Aldrich) was injected intraperitoneally in the middle of a behavioral session in trained animals, and the same amount of saline was injected in the same animals on another day as a control. Note that 0.3 mg/kg is considered a fairly high dose in previous work[14-16]. Neural activity and mice performance in BMI task one hour before and one hour after drug injection were monitored and analyzed for comparison.

Lever-Pressing Task

RGS9L-Cre/Nr1$^{f/f}$ animals and their littermate controls were first trained to acquire a regular sequence task under fixed-ratio eight (FR8) schedule for one week as described before[5]. After that, a differential reinforcement sequence training schedule was introduced where only eight lever presses performed within 16 seconds will be reinforced. Self-paced lever press sequences were defined by either the statistics of lever pressing for each animal (bimodal or Poisson distribution) or by a bout of licks interrupting lever pressing. The sequence length and duration were thus calculated based on each individual sequence, and the within-sequence press rate was computed by the ratio of sequence length (≥ 2 presses) and the corresponding sequence duration[5].

References

1. Dang M.T. *et al.* Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 15254-15259 (2006).
2. Zhong, Y. & Bellamkonda, R.V. Dexamethasone-coated neural probes elicit attenuated inflammatory response and neuronal loss compared to uncoated neural probes. *Brain Res.*, **1148**, 15-27 (2007).
3. Han, Y.K., Köver, H., Insanally, M.N., Semerdjian, J.H., & Bao, S. Early experience impairs perceptual discrimination. *Nature Neurosci.*, **10**, 1191-1197 (2007).
4. Yokota, T., Saito, Y., & Miyatake, T. Conduction slowing without conduction block of compound muscle and nerve action potentials due to sodium channel block. *J. Neurol. Sci*., **124**, 220-224 (1994).

5. Jin, X. & Costa, R.M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, **466**, 457-462 (2010).
6. Belova, M.A., Paton, J.J., Morrison, S.E., & Salzman, C.D. Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron*, **55**, 970-984 (2007).
7. Venkatraman, S., Jin, X., Costa, R.M., & Carmena, J.M. Investigating neural correlates of behavior in freely behaving rodents using inertial sensors. *J. Neurophys*., **104**, 569-575 (2010).
8. Jarvis, M. & Mitra, P. Sampling properties of the spectrum and coherency of sequences of action potentials. *Neural Comput.*, **13**, 717-749 (2001).
9. Thomson, D. Spectrum estimation and harmonic analysis. *Proc IEEE*, **70**, 1055-1096 (1982).
10. Mitra, P. & Pesaran, B. Analysis of dynamic brain imaging data. *Biophys. J.*, **76**, 691-708 (1999).
11. Zeitler, M., Fries, P., & Gielen, S. Assessing neuronal coherence with single-unit, multi-unit, and local field potentials. *Neural Comp.*, **18**, 2256-2281 (2006).
12. Schoffelen, J.M., Poort, J., Oostenveld, R., & Fries, P. Selective movement preparation is subserved by selective increases in corticomuscular gamma-band coherence. *J. Neurosci.*, **31**, 6750-6758 (2011).
13. Suner, S., Fellows, M.R., Vargas-Irwin, C., Nakata, G.K., & Donoghue, J.P. Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex. *IEEE TNSRE*, **13**, 524-541 (2005).
14. Butelman, E.R. A novel NMDA antagonist, MK-801, impairs performance in a hippocampal-dependent spatial learning task. *Pharmacol. Biochem. Behav.*, **34**, 13-16 (1989).

15. Lee, J.L., Milton, A.L., & Everitt, B.J. Reconsolidation and extinction of conditioned fear: inhibition and potentiation. *J. Neurosci.*, **26**, 10051-10056 (2006).

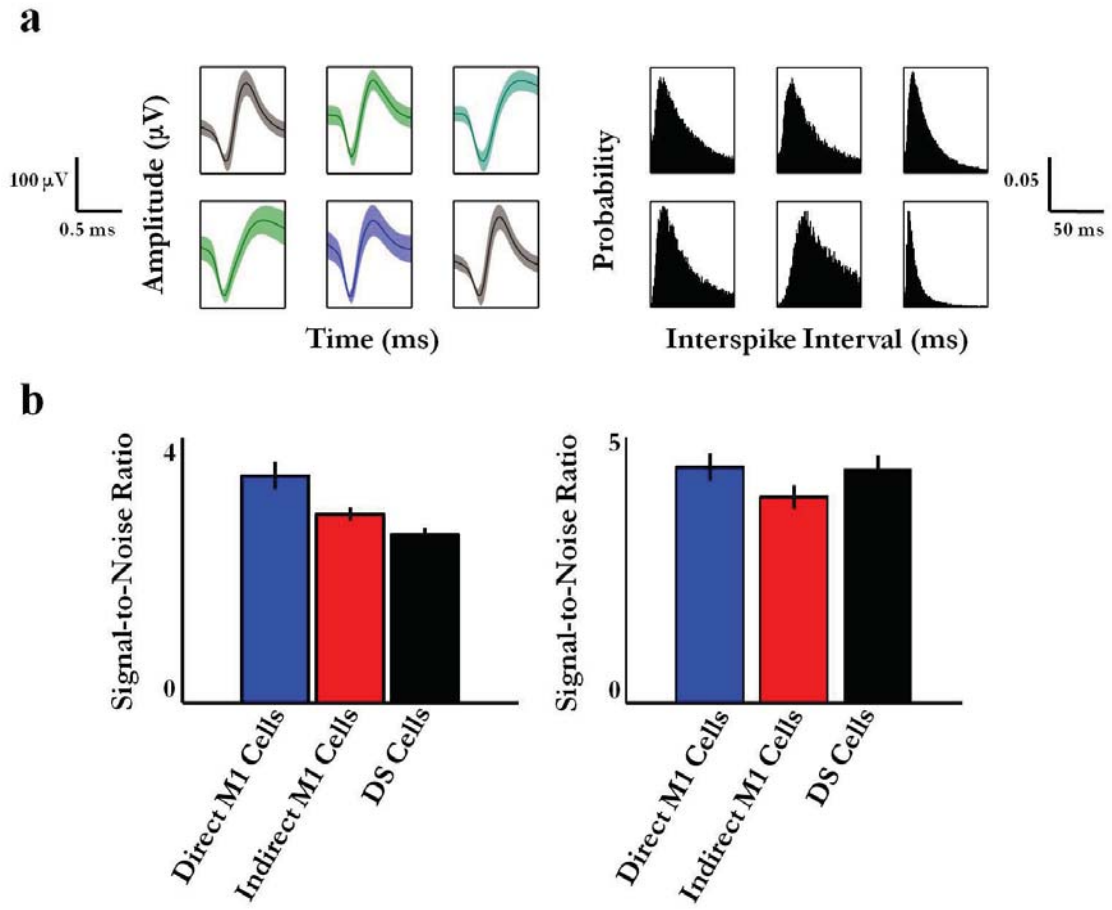
16. Autry, A.E., Adachi, M., Nosyreva, E., Na, E.S., Los, M.F., Cheng, P.F., Kavalali, E.T., & Monteggia, L.M. NMDA receptor blockade at rest triggers rapid behavioural antidepressant responses. *Nature*, **475**, 91-95 (2011).

SUPPLEMENTARY VIDEOS

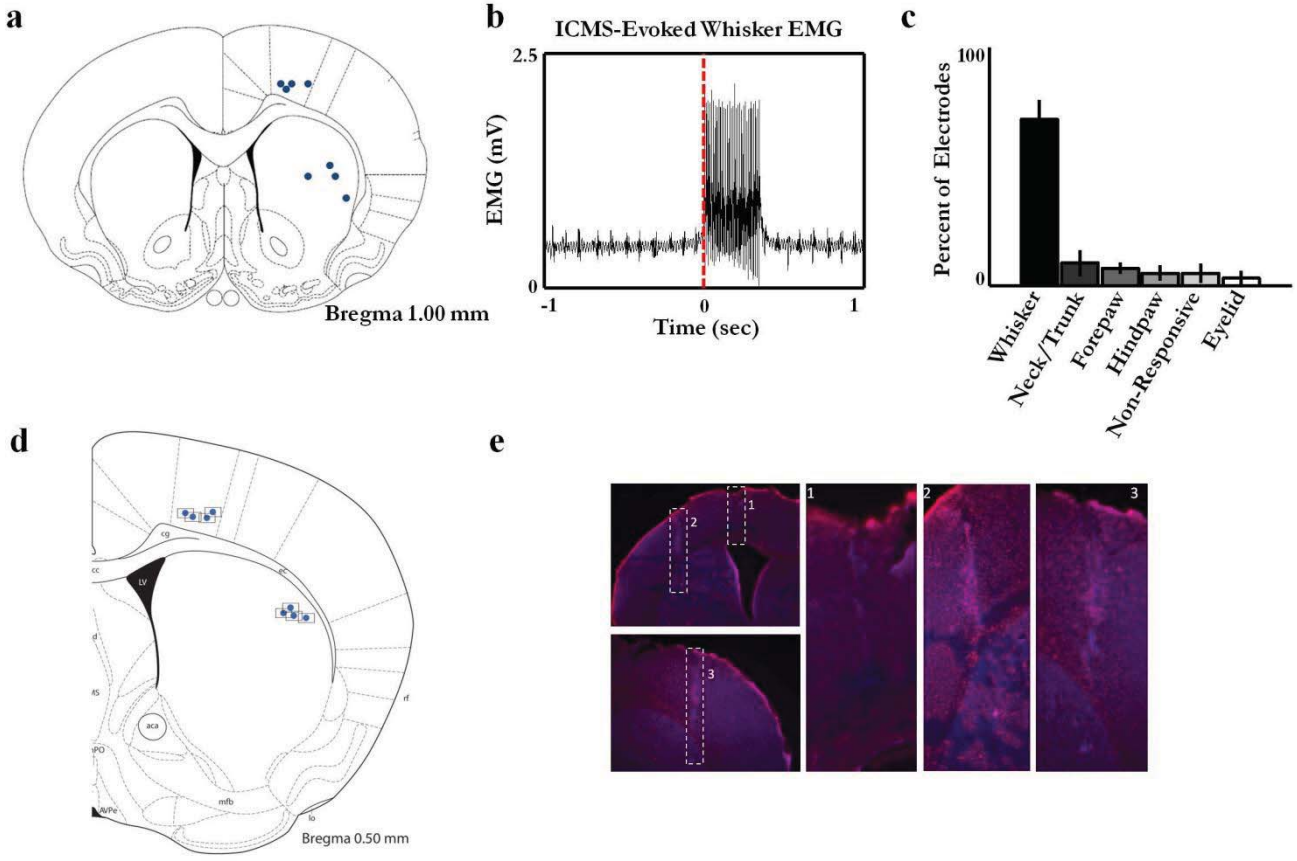
Supplementary Video 1: Example of locomotor activity for a control mouse with electrode implantation and used in the BMI experiments.

Supplementary Video 2: Example of locomotor activity for a RGS9L-Cre/Nr1$^{f/f}$mouse with electrode implantation and used in the BMI experiments.

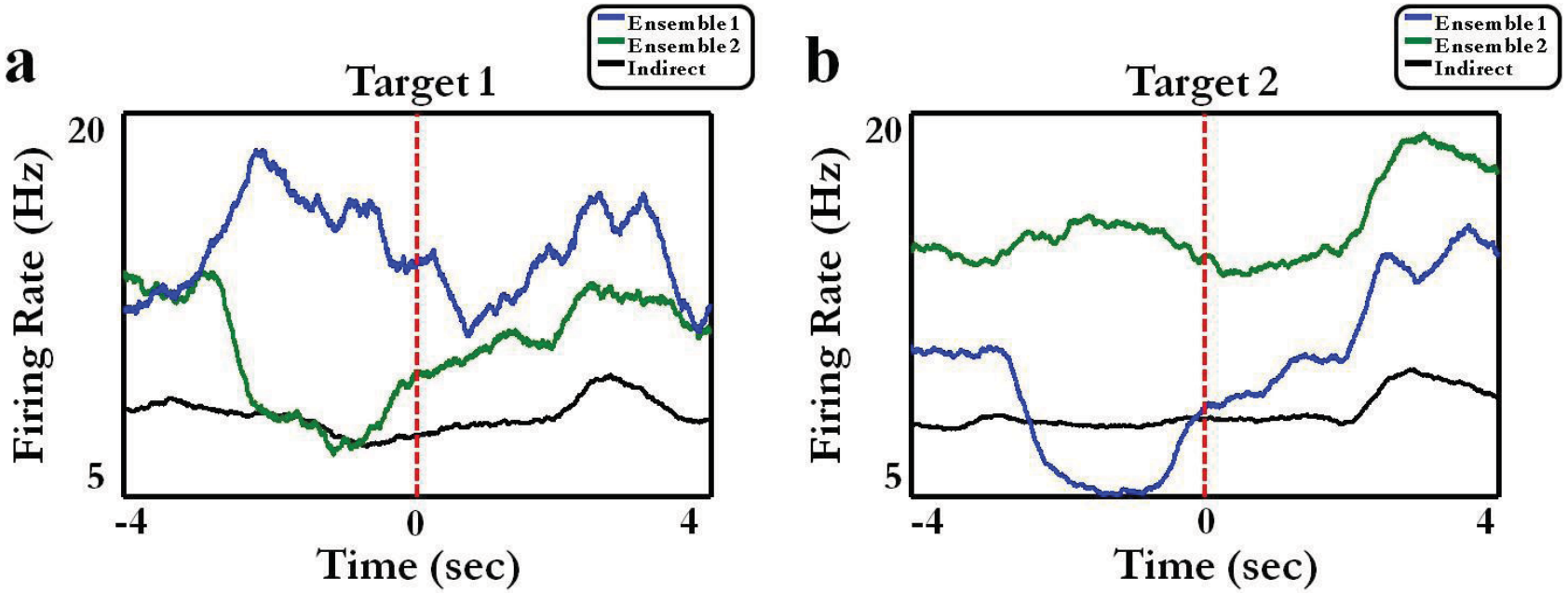
SUPPLEMENTARY FIGURES



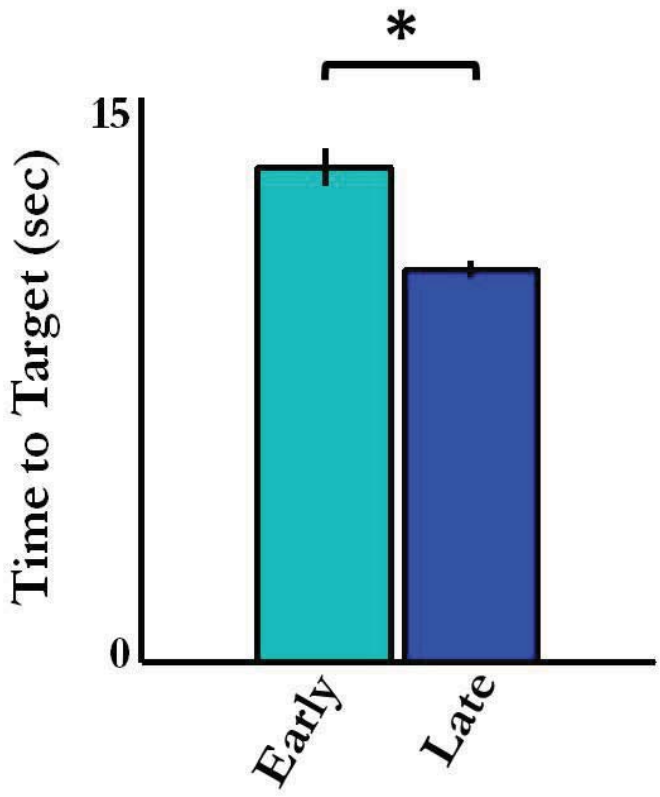
Supplementary Figure 1. Simultaneous recording from M1 and the DS. a. Representative waveforms (left) and interspike interval distributions (right) recorded during the experiment. Waveforms were chosen based on waveform reproducibility, signal strength, and the refractory period. **b.** Quantification of the signal-to-noise ratio of our neural recordings for M1 cells entered into the transform algorithm ("Direct M1 Cells"), M1 cells not entered into the transform algorithm ("Indirect M1 Cells"), and DS cells in rats (left) and mice (right).



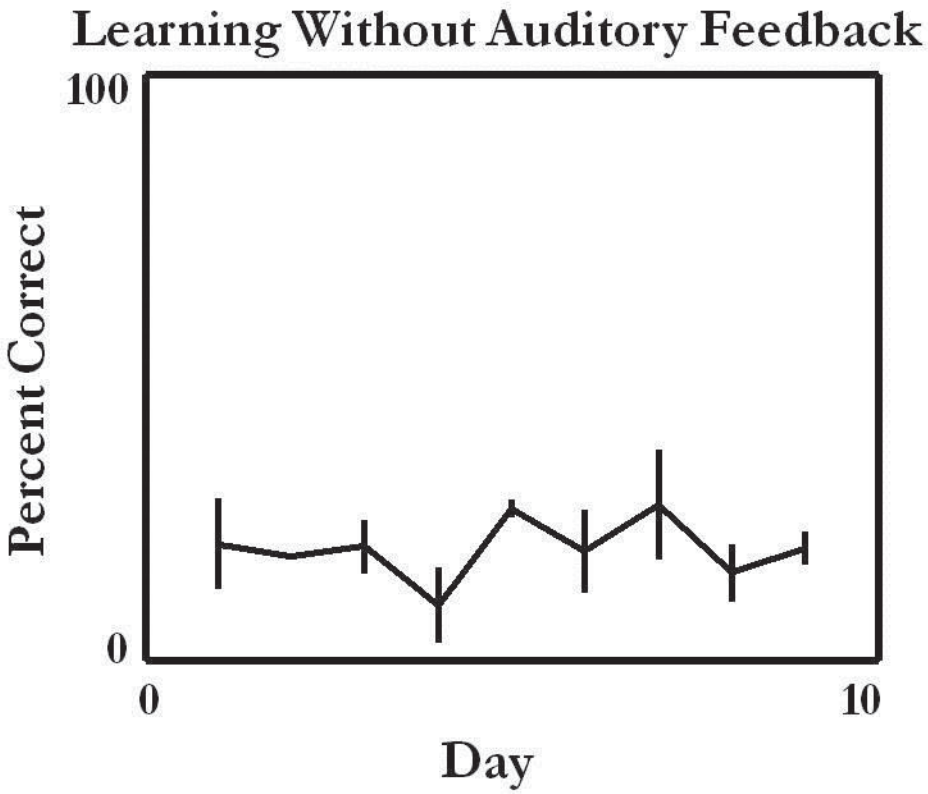
Supplementary Figure 2. Verification of Electrode Placement. a. Results of histology for 4 rats showing placement of electrode tips in both M1 and the DS. **b.** Representative example of EMG recorded from the intrinsic muscles of the mystacial pad following ICMS applied through our recording electrodes. **c.** Functional verification of M1 recording sites showing the percentage of electrodes that evoked motor twitches in each body part following ICMS application to the recording sites. **d.** Results of histology in mice showing placement of electrode tips in both M1 and the DS. **e.** Example of electrode tracks seen following staining with DAPI and propidium iodide.



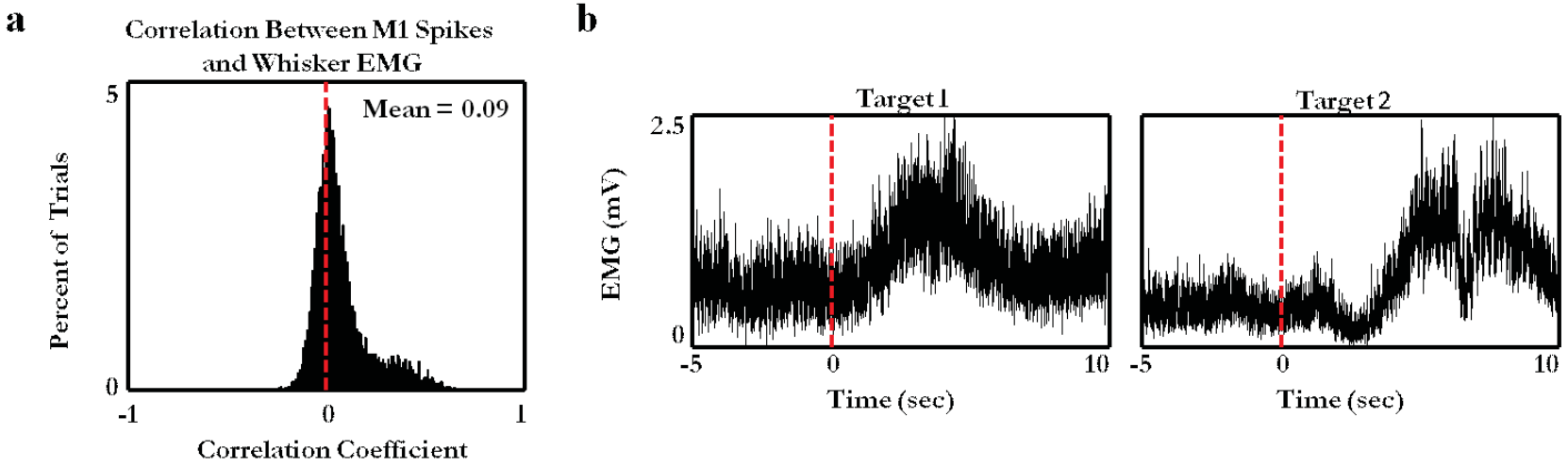
Supplementary Figure 3. Mean M1 firing rates across all animals. M1 firing rates for units in ensemble 1 (blue), ensemble 2 (green), and units not used for the task (black) averaged across all animals and time-locked to achievement of target 1 (a) or target 2 (b). Rodents were producing the desired ensemble rate modulations to properly perform the task. The units not used for the task did not show modulation during target reaching (before 0) but did show modulation during movement to retrieve the earned reward (between +2 and +4 secs, see also accelerometer data).



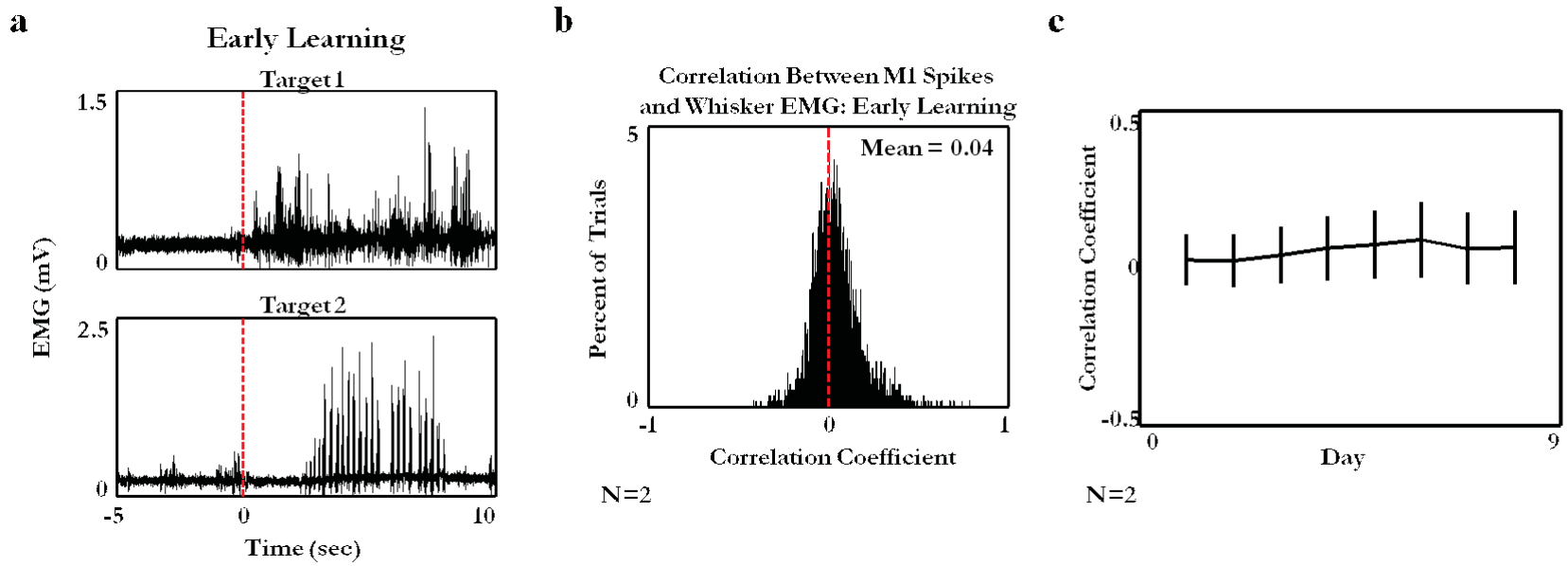
Supplementary Figure 4. Reduction in the mean time to target during learning. In addition to an increase in the percentage of correct trials, the time from trial initiation to target achievement decreased significantly from early to late in learning ($p < 0.001$).



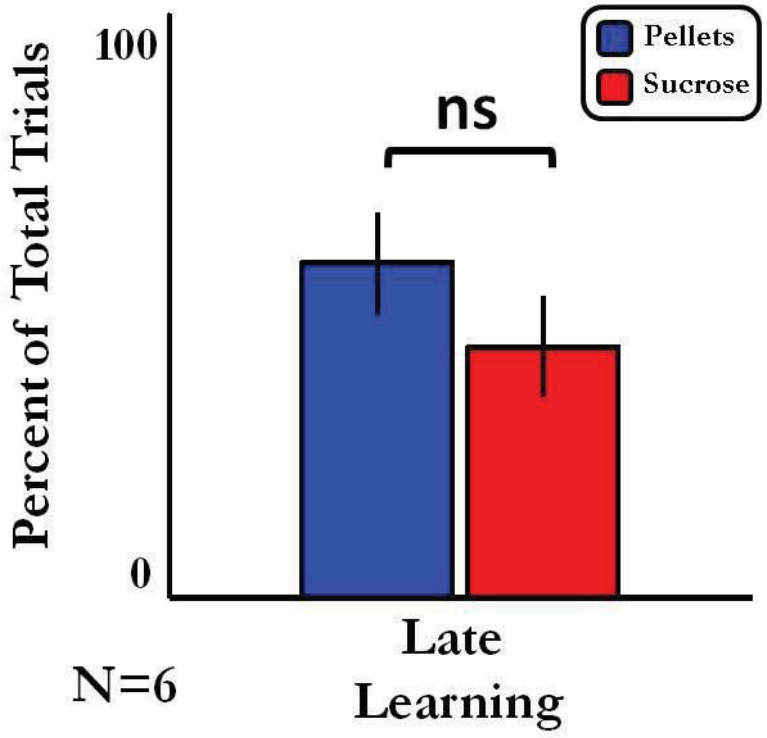
Supplementary Figure 5. Learning without auditory feedback. When auditory feedback about performance was not supplied to rats, the percentage of correct trials did not increase significantly over the course of 9 days of training.



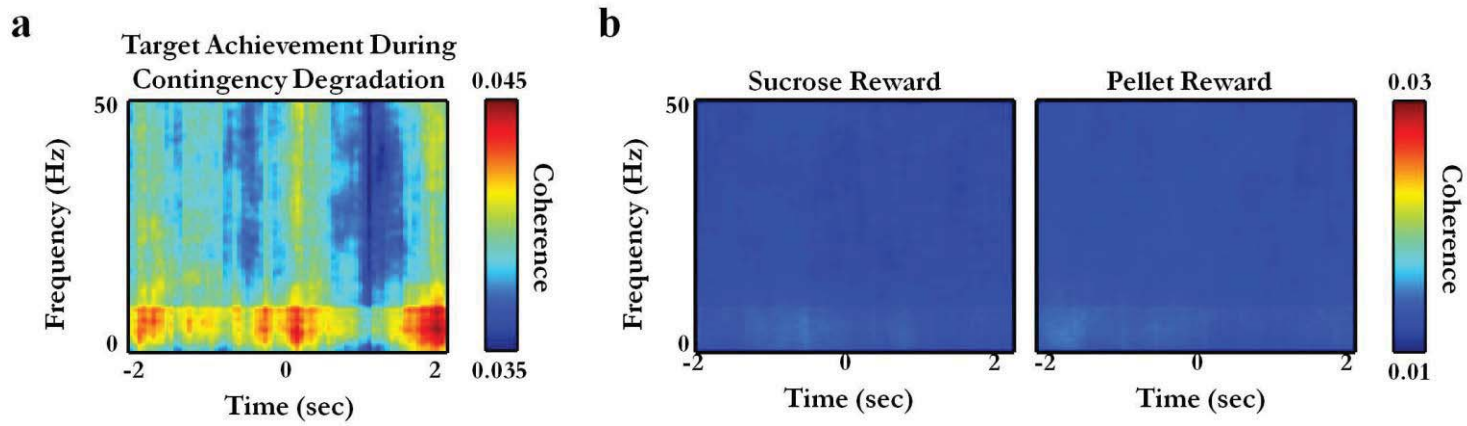
Supplementary Figure 6. EMG from the mystacial pad during task performance. a. The distribution of correlation coefficients between M1 spiking activity and the EMG of the mystacial pad is not significantly different from zero, with values a mean value of 0.092 ± 0.003. **b.** Mean EMG activity across a full behavioral session in 2 animals time-locked to reward delivery after animals hit target 1 (left) or target 2 (right). EMG activity is very low before target achievement, but high afterwards as animals retrieve reward.



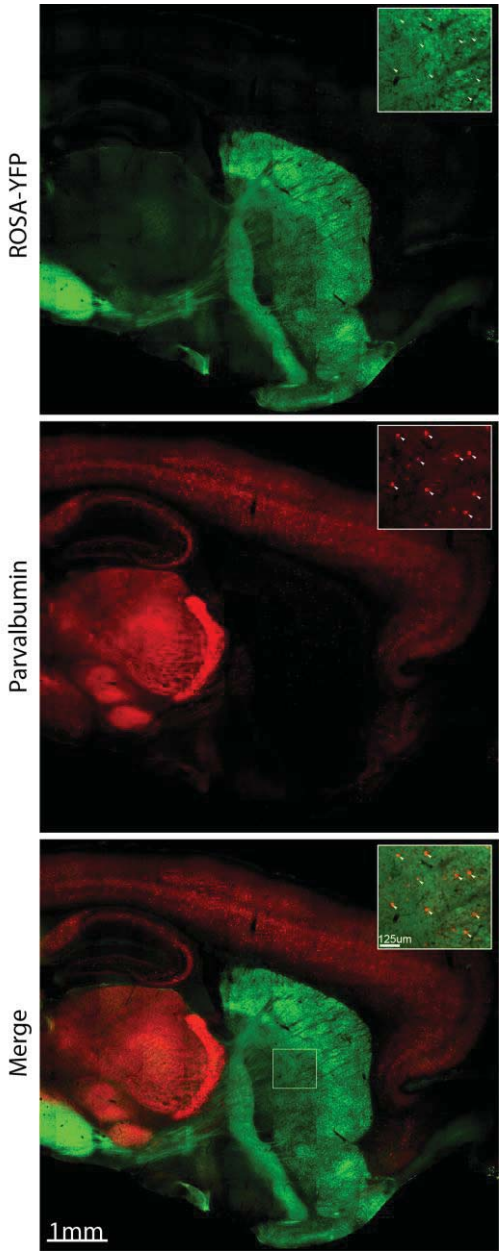
Supplementary Figure 7. Rats do not show changes in EMG activity during target reaching throughout learning. a. In early learning, there is no strong EMG signal before target achievement, despite clear activity afterwards as animals retrieve and consume reward. **b.** The distribution of correlation coefficients between EMG activity and M1 spiking activity for all trials in a session in early learning is not significantly different from zero. **c.** Correlation coefficients between EMG activity and M1 spiking activity across 8 days of learning suggest that animals do not rely on physical movements at any stage of training.



Supplementary Figure 8. No reward preference in late learning. Animals did not respond significantly more for either pellets or sucrose reward in late learning, validating that response differences seen during the devaluation test were not due to differences in baseline preference.



Supplementary Figure 9. Coherence remains high during contingency degradation. a. When target achievement was no longer rewarded, coherence between M1 spikes and DS spikes remained high surrounding target achievement. **b.** Coherence was very low when rewards were delivered to the animal during the contingency degradation experiment. Together, these suggest that coherence is not due to reward expectation, but rather relates to task performance.



Supplementary Figure 10. RGS9L-Cre deletes in both striatonigral and striatopallidal medium spiny neurons, but does not delete in parvalbumin-positive interneurons in striatum.

Top panel. RGS9L-cre mice were crossed with ROSA-YFP reporter mice to determine where Cre deletion occurs in this line (n=4). Cre-driven recombination occurs very prominently in striatum, and is largely absent in other areas like cortex and hippocampus. The staining of cell bodies and axons confirms that Cre deletion occurs in both striatonigral and striatopalidal medium spiny neurons (direct and indirect pathway respectively), as can be attested by the prominent labeling of axonal projections in *globus pallidus externus, globus pallidus internus* and *substantia nigra pars reticulata* (in this order, form right to left). Inset shows a higher magnification of striatum showing that recombination occurs in most neurons.

Middle panel.

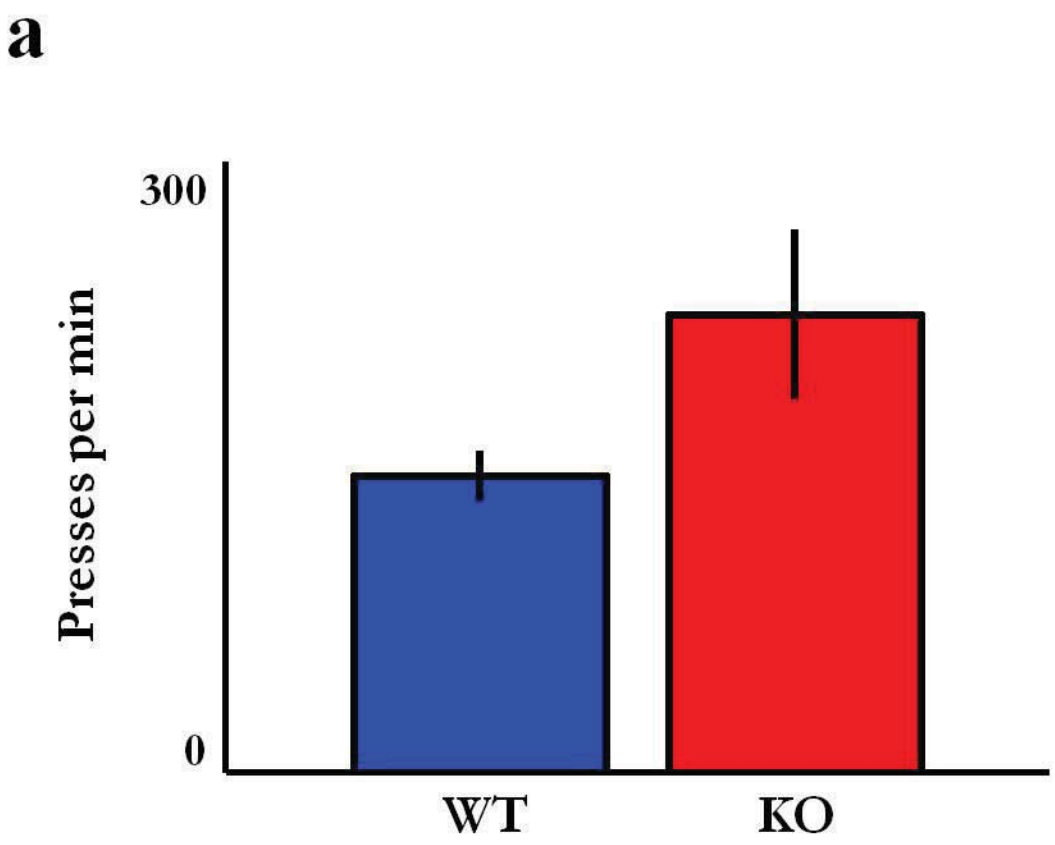
Immunohistochemistry against parvalbumin (red) reveals a different pattern of expression of PV and RGS9L-Cre, with parvalbumin being expressed heavily in cortex, hippocampus and thalamus and in parvalbumin-positive striatal interneurons. Inset shows Parvalbumin-positive interneurons in striatum marked with a white arrow.

Bottom panel.

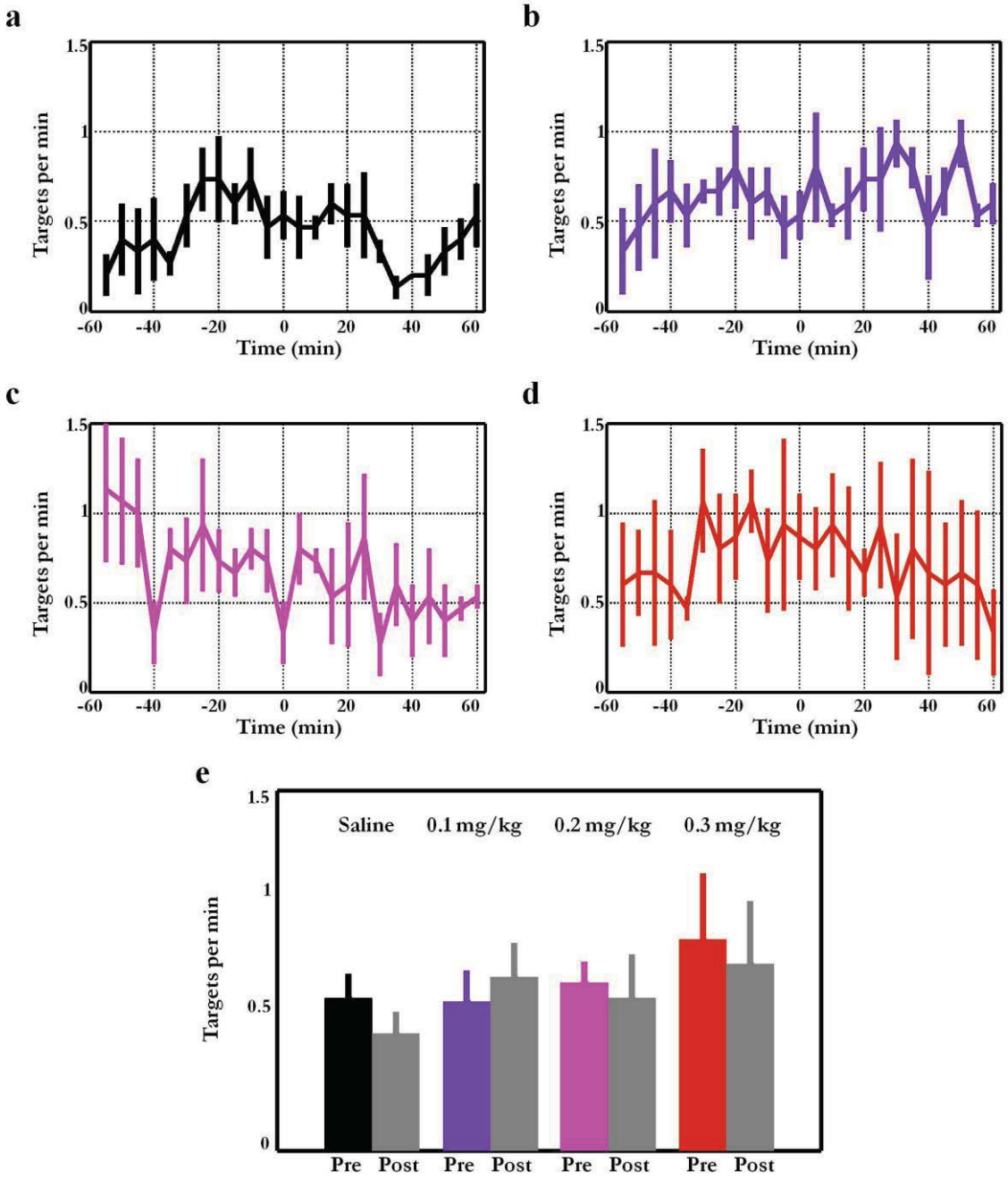
Merge of the two pictures above showing no co-localization between the YFP (RGS9L-Cre) and parvalbimun. Inset shows that there is no cre-dependent recombination in parvalbumin positive interneurons in striatum.

Note: Three different lines with selective deletion of NMDA receptors in striatum by crossing *Cre* mice with *NMDAR1-loxP* mice were published to date. Mice resulting from the cross between *NMDAR1-loxP* and *Dlx5/6-Cre* display a severe phenotype with reduced food intake, retarded growth starting and eventually dye at around postnatal age 20 (1). Lines where the *NMDAR1-loxP* mice were crossed with *RGS9L-Cre* and *GPR88-Cre* show no lethality, but still lack NMDA currents in most medium spiny neurons, and show no corticostriatal LTP (2,3). The differences between the lines have been attributed to a variety of factors including the extent of deletion in striatum (*Dlx5/6*-Cre expresses in all types of striatal neurons including interneurons while *GPR88-Cre* mice do not target interneurons and the *RGS9L-Cre* line spares at least parvalbimin-positive interneurons, Supplementary Figure 10), the specificity of deletion in areas other than striatum, and the age at which the different Cre lines start deleting (*Dlx5/6*-Cre is the earliest deleting from embryonic day 12.5). We used *NMDAR1-loxP* mice crossed with the *RGS9L-Cre* in the current study because the mice survive and move normally, can learn operant tasks and execute fast motor sequences, but still lack corticostriatal LTP and are impaired in sequence learning and natural skill learning (2, 4).

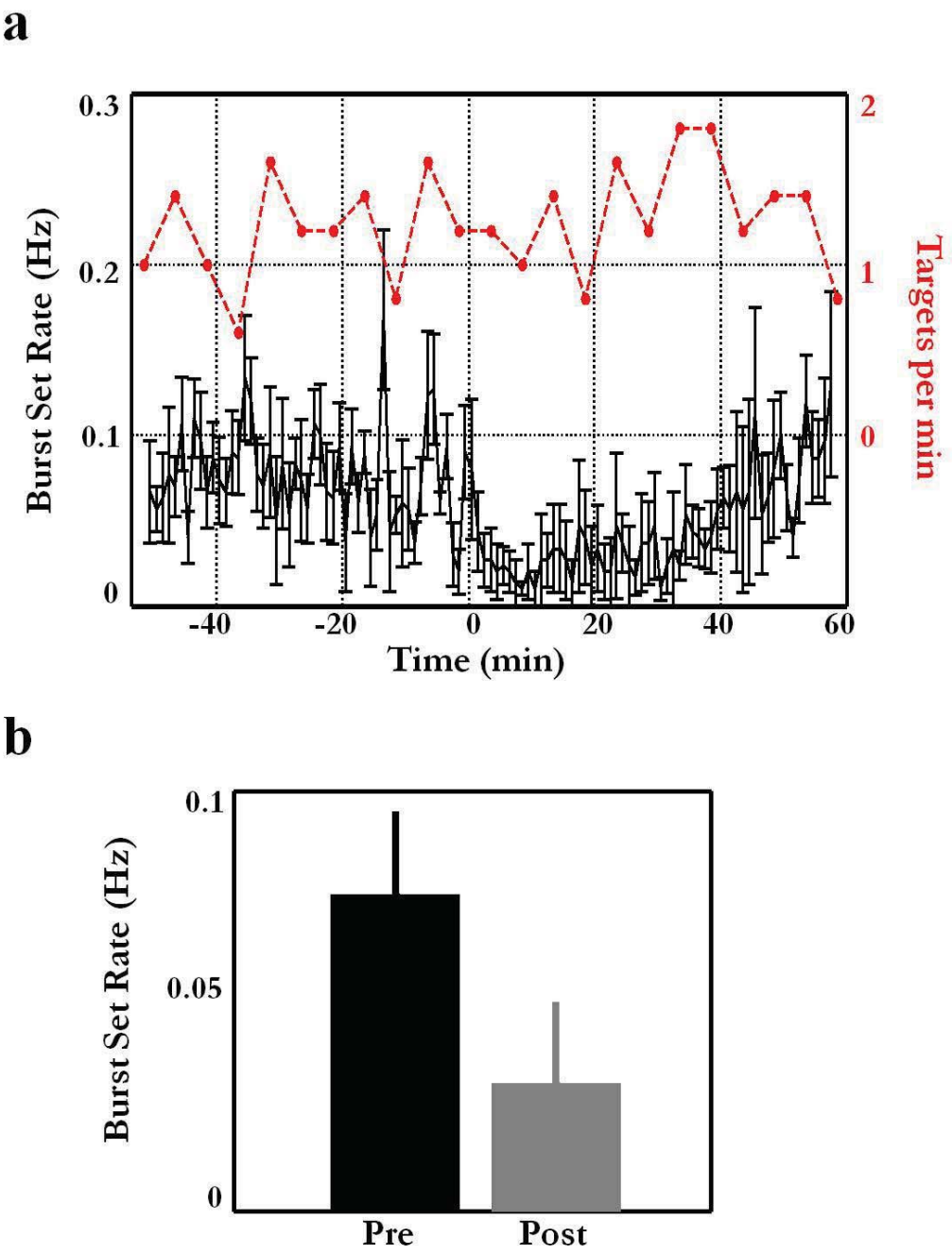
1- Functional disturbances in the striatum by region-specific ablation of NMDA receptors. Ohtsuka N, Tansky MF, Kuang H, Kourrich S, Thomas MJ, Rubenstein JL, Ekker M, Leeman SE, Tsien JZ. Proc Natl Acad Sci U S A. 2008 Sep 2;105(35):12961-6.
2- Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum.Dang MT, Yokoi F, Yin HH, Lovinger DM, Wang Y, Li Y.Proc Natl Acad Sci U S A. 2006 Oct 10;103(41):15254-9.
3- Severely Impaired Learning and Altered Neuronal Morphology in Mice Lacking NMDA Receptors in Medium Spiny Neurons. Beutler LR, Eldred KC, Quintana A, Keene CD, Rose SE, Postupna N, Montine TJ, Palmiter RD. PLoS One. 2011;6(11):e28168
4- Start/stop signals emerge in nigrostriatal circuits during sequence learning. Jin X, Costa RM. Nature. 2010 Jul 22;466(7305):457-62.



Supplementary Figure 11. Within-sequence press rates of the RGS9L-Cre/Nr1$^{f/f}$**and littermate controls under the differential reinforcement schedule FR8/16s.** Consistent with previous work [4], RGS9L-Cre/Nr1$^{f/f}$showed higher within-sequence press rate than littermate controls ($p < 0.05$), indicating they have no impairment in motor speed per se.



Supplementary Figure 12. Blockage of NMDA receptors after training has no effect on BMI performance. a-d. Behavioral performance in trained mice before and after systemic administration (i.p.) of NMDA receptor antagonist MK-801 (a: saline, b: 0.1 mg/kg; c: 0.2 mg/kg; d: 0.3 mg/kg). Time zero in each case indicates time of injection **e.** Summary of the effect of NMDA receptor blockade on task performance in trained animals. There is no general effect of drug treatment ($F_{2,7} = 0.45$, $P = 0.86$; $P > 0.05$ for all pairs of comparison).



Supplementary Figure 13. Blockage of NMDA receptors impairs striatal MSNs burst firing. a. The burst activity of striatal MSNs (black line) was greatly decreased after application of NMDA antagonist MK-801 at higher dose (0.3 mg/kg at time zero), but the mouse was able to maintain its performance throughout the time (red line). **b.** Statistical results of striatal MSNs burst activity pre- and post- NMDA receptor blockage. There was a significant decrease of burst set rate ($P < 0.01$).