Articles in PresS. J Neurophysiol (December 28, 2016). doi:10.1152/jn.00423.2016

Operant conditioning of neural activity in freely behaving monkeys with intracranial reinforcement

Ryan W. Eaton[1], Tyler Libey[2,3], Eberhard E. Fetz[1,2,3]

1. Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195-7290
2. Department of Bioengineering, University of Washington, Seattle, WA 98195- 5061
3. Center for Sensorimotor Neural Engineering, NSF ERC, University of Washington, Seattle, WA 98195- 6271

Running title: Conditioning neural activity with intracranial reinforcement

Corresponding Author:
Eberhard E. Fetz
Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195-7290
e-mail: fetz@uw.edu

Abstract

Operant conditioning of neural activity has typically been performed under controlled behavioral conditions using food reinforcement. This has limited the duration and behavioral context for neural conditioning. To reward cell activity in unconstrained primates, we sought sites in nucleus accumbens (NAc) whose stimulation reinforced operant responding. In three monkeys NAc stimulation sustained performance of a manual target-tracking task, with response rates that increased monotonically with increasing NAc stimulation. We recorded activity of single motor cortex neurons and documented their modulation with wrist force. We conditioned increased firing rates with the monkey seated in the training booth and during free behavior in the cage using an autonomous head-fixed recording and stimulating system. Spikes occurring above baseline rates triggered single or multiple electrical pulses to the reinforcement site. Such rate-contingent, unit-triggered stimulation was made available for periods of 1-3 minutes separated by 3-10 minute time-out periods. Feedback was presented as event-triggered clicks both in-cage and in-booth, and visual cues were provided in many in-booth sessions. In-booth conditioning produced increases in single neuron firing probability with intracranial reinforcement in 48 of 58 cells. Reinforced cell activity could rise > 5 times that of non-reinforced activity. In-cage conditioning produced significant increases in 21 of 33 sessions. In-cage rate changes peaked later and lasted longer than in-booth changes but were often comparatively smaller, between 13 and 18 percent above non-reinforced activity. Thus intracranial stimulation reinforced volitional increases in cortical firing rates during both free behavior and a controlled environment, although changes in the latter were more robust.

New and Noteworthy

Closed-loop brain-computer interfaces (BCI) were used to operantly condition increases in muscle and neural activity in monkeys by delivering activity-dependent stimuli to an intracranial reinforcement site (nucleus accumbens). We conditioned increased firing rates with the monkeys

seated in a training booth and also, for the first time, during free behavior in a cage using an autonomous head-fixed BCI.

Introduction

Volitional control of neural activity is critical for reliable and robust control of brain-machine interfaces (BMI). Indeed, BMIs can be seen as a form of neurofeedback that allows the user to see the consequences of neural activity and change that activity to optimize control of the external device (Fetz, 2007). However, BMI control is only a subset of the possible range of volitional control of neural activity that can be explored directly with operant conditioning. Traditional techniques for operant conditioning of behavior in monkeys have limited the scope of investigation to specific tasks, using food reward and visual feedback delivered in a training booth. Constrained, task-related movements differ from natural behavior, and correlations between neural activity and movement established under particular task conditions may not hold under non-task conditions (Caminiti et al., 1990) (Aflalo and Graziano, 2006) (Jackson et al., 2007). The vast majority of nonhuman primate research involving trained behavior has employed rewards in the form of food or water (Taylor et al., 2002, Carmena et al., 2003, Jackson et al., 2006), further limiting the circumstances in which neural activity was explored. Here we present a novel mechanism for rewarding neural activity during natural behavior using a closed-loop system delivering neurally contingent brain stimulation reward (BSR).

Olds & Milner (Olds and Milner, 1954) (Olds, 1958) demonstrated that rats would press bars and navigate mazes for BSR, which could reinforce operant responding as effectively as more conventional food and liquid rewards. Later work by David Hiatt attempted to condition increases in single-unit activity using burst-triggered BSR in rats (Hiatt, 1972). As candidates for conditioning he sought cells in hippocampus, cerebellum, midbrain and superior colliculus that were not movement related. Recently, BSR was used to elicit rate increases in prefrontal cortex neurons of freely behaving rats (Widge and Moritz, 2014). The ability of freely moving rats to differentially control small groups of cortical neurons was demonstrated with food reward and continuous

auditory feedback (Koralek et al., 2012).

Several studies have explored the efficacy of BSR in non-human primates. In a freely behaving chimpanzee Delgado et al deployed wireless closed-loop stimulation of reticular formation sites contingent on oscillations in amygdala field potentials (Delgado et al., 1970). The triggering neural oscillations disappeared after a day of activity-dependent stimulation, indicating that this form of stimulation was aversive. Later work showed that monkeys will perform simple bar-press tasks for BSR in several structures, including the orbitofrontal cortex, lateral hypothalamus, amygdala, medio-dorsal nucleus of the thalamus and nucleus accumbens (Briese and Olds, 1964, Routtenberg et al., 1971, Rolls et al., 1980, Bichot et al., 2011, Bowden DM, 2015).

An interesting open question is whether monkeys can learn to control activity of single neurons with intracranial electrical stimulation as the sole source of reinforcement. This would allow operant conditioning to be performed during prolonged periods of free behavior, providing extended time and behavioral range to learn volitional control of neural response patterns. BSR would enable delivery of reinforcement that is temporally more precise than food or water rewards, and less disruptive of ongoing behavior. In this study we sought to operantly condition activity of motor cortex neurons and electromyographic (EMG) activity of proximal limb muscles, using activity-contingent BSR at sites confirmed to sustain behavior in a target-tracking task. To compare the effects of the environment, we conditioned these activities both in the training booth and as the monkeys moved freely about their home cage.

Methods

Subjects and Training

We used 3 male Macaca Nemestrina monkeys P, D and J (4-6 years old, weight 6.0, 5.6 and 4.0 kg). All surgical, training and handling procedures were approved by the University of Washington Institutional Animal Care and Use Committee.

Prior to surgeries, monkeys were trained to perform a 1D center-out force-target-tracking (FTT) task in which

isometric wrist torque controlled the position of a cursor on a screen. When the cursor entered a target and remained inside for the required time (1 second or less) a fruit sauce reward signaled completion of the trial. Target placement on the screen determined the required direction and magnitude of flexion or extension torque about the wrist. Peripheral targets were presented in random order with equal probability. Training was complete when monkeys moved directly from center to each target, and held it inside for at least 1 second. During experiments, the FTT task was performed daily to elicit task-related cell firing in motor cortex.

Surgery and Implantation

Cranial microwires and arrays of up to 16 cannulae were implanted in each monkey. The microwire arrays (Jackson and Fetz, 2007) were positioned to advance along layer V in the caudal bank of the precentral gyrus, where somata of many force-correlated cells (including corticomotoneuronal (CM) cells) have been identified (Rathelot and Strick, 2009), (Smith and Fetz, 2009). The cannulae were positioned stereotaxically to guide subsequent stimulating electrodes to the nucleus accumbens (NAc). Cannula-length stylets were placed in all guide tubes and the protruding surface of the array was sealed in silastic. The open space between craniotomy and array was packed with antibiotic-infused gelfoam. An acrylic base around the implantation site and surrounding cranial screws formed the base for a cylindrical titanium chamber enclosing the microwire cannula arrays and Neurochip (Zanos et al., 2011). Rhodes SNEX-100 concentric bipolar electrodes were inserted subsequently into the cannulae after cold-sterilization of the chamber interior and electrodes with cidex.

To identify potential intracranial reinforcement electrode implant sites, we co-registered a magnetic resonance image (MRI) and digitized brain atlas data (**National Primate Research Center**, 1991-present) to determine the stereotaxic coordinates of prospective midbrain reinforcement loci (Figure 1). Monkey P underwent MRI scanning prior to surgical implantation. Monkeys D and J were of similar size as atlas subjects, so MRIs were not deemed necessary. We selected coronal image slices located +3 mm rostral from the anterior commissure that contained the largest cross-section of the Nucleus Accumbens (NAc). Stereotaxic coordinates of the target locus were measured relative to medial-lateral center and ear-bar zero. A straight-line diagonal

path to the target locus (center of NAc) that was 15 degrees lateral right with respect to the dorso-ventral axis in the right hemisphere avoided major blood vessels and regions governing autonomic function. To address the possibility of positioning error of entry sites, we implanted an array of 16 parallel cannulae spaced 1 to 1.5 mm apart in a 10×10 mm grid centered at the best point of entry. Thus, in cases of slight angle misalignment or entry location, the target locus might still be reachable by an electrode inserted in one of the neighboring cannulae. Following implantation, unused cannulae were occluded with stylets and sealed with silastic to block potential cranial infection.

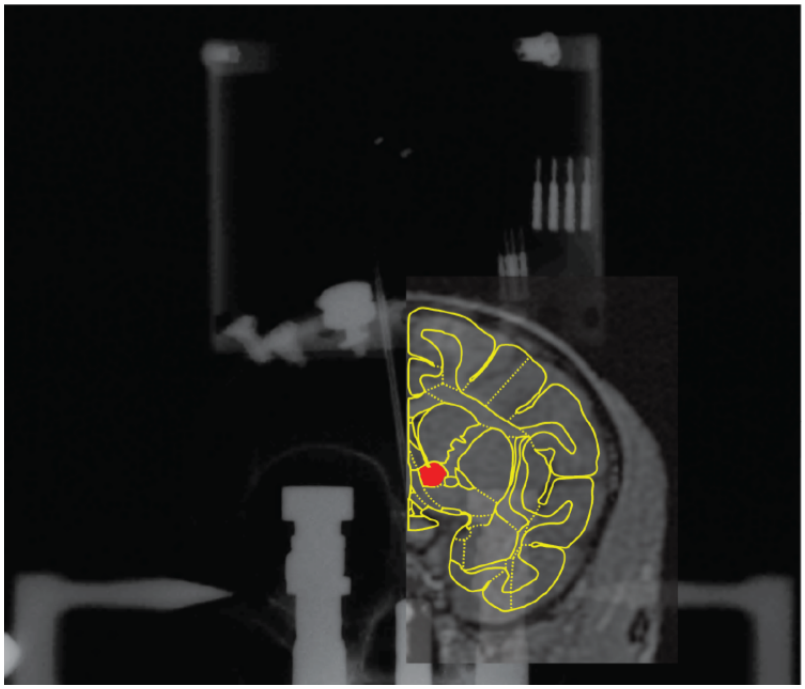


Figure 1. Co-registration of cranial X-ray, MRI and brain atlas images. A coronal map of the M. Nemestrina brain was morphed and superimposed on an MRI of Monkey P's brain. Both are positioned over an X-ray image showing the chamber and housed hardware. NAc shown in red.

In monkey P, in addition to cranial implant procedures, we implanted pairs of EMG wires in three proximal muscles of the monkey's right arm: the biceps brachii, triceps brachii and lateral deltoid. Muscle activity was first operantly conditioned to verify efficacy of BSR in free behavior. The EMG wires were routed subcutaneously around the shoulder, up the back and neck and terminated in connectors located inside the cranial chamber for signal processing by the Neurochip.

Verification of brain stimulation reward

To identify intracranial brain sites whose stimulation sustains operant responding, we compared response rates occurring during reinforcement (R) and visual feedback-only (FO) blocks in a FTT task. During R blocks, each completed flexion or extension target hold triggered BSR. In FO blocks no stimulation was delivered, regardless of task performance, but the FTT task could be performed. R and FO blocks were interleaved with non-reinforcing (NR) blocks in which neither feedback nor reward were available. Stimulation consisted of trains of symmetric biphasic square-wave current pulses. A low-frequency tone during R blocks served as a discriminatory stimulus (in addition to FTT task auditory cues for target acquisition). Candidate sites were considered to be "positively reinforcing" when monkeys performed wrist FTT at significantly greater rate during R blocks than during FO blocks.

Rate-contingent spike-triggered (RCST) stimulation

Validated BSR sites were used to operantly condition cortical cell and muscle activity in two different settings: a traditional in-booth setting using rack mounted equipment for recording and stimulation and an in-cage setting using the Neurochip system (Figure 2A, B). The Neurochip (NC) employs an autonomous, battery-powered computer chip programmed to detect and reward cell and muscle activity while monkeys moved freely about their cages (Mavoori et al., 2005). It discriminated cortical cell or EMG activity patterns using dual time-amplitude window discrimination and delivered stimuli contingent on discriminated events in real time. The Neurochip2- HV is a second-generation version with improved capabilities for storage, processing and stimulus range (Zanos et al., 2011). Alternating R/NR reinforcement schedules were used to distinguish the effects of BSR in the operant conditioning paradigm. FO blocks were not used during these experiments. The in-booth experiments utilized audio and visual feedback to distinguish between the periods, whereas the in-cage experiments relied solely on audio feedback. The in-booth experiments lasted between 1 and 6 hours, while the in-cage free-behavior sessions lasted considerably longer: 3-20 hours.

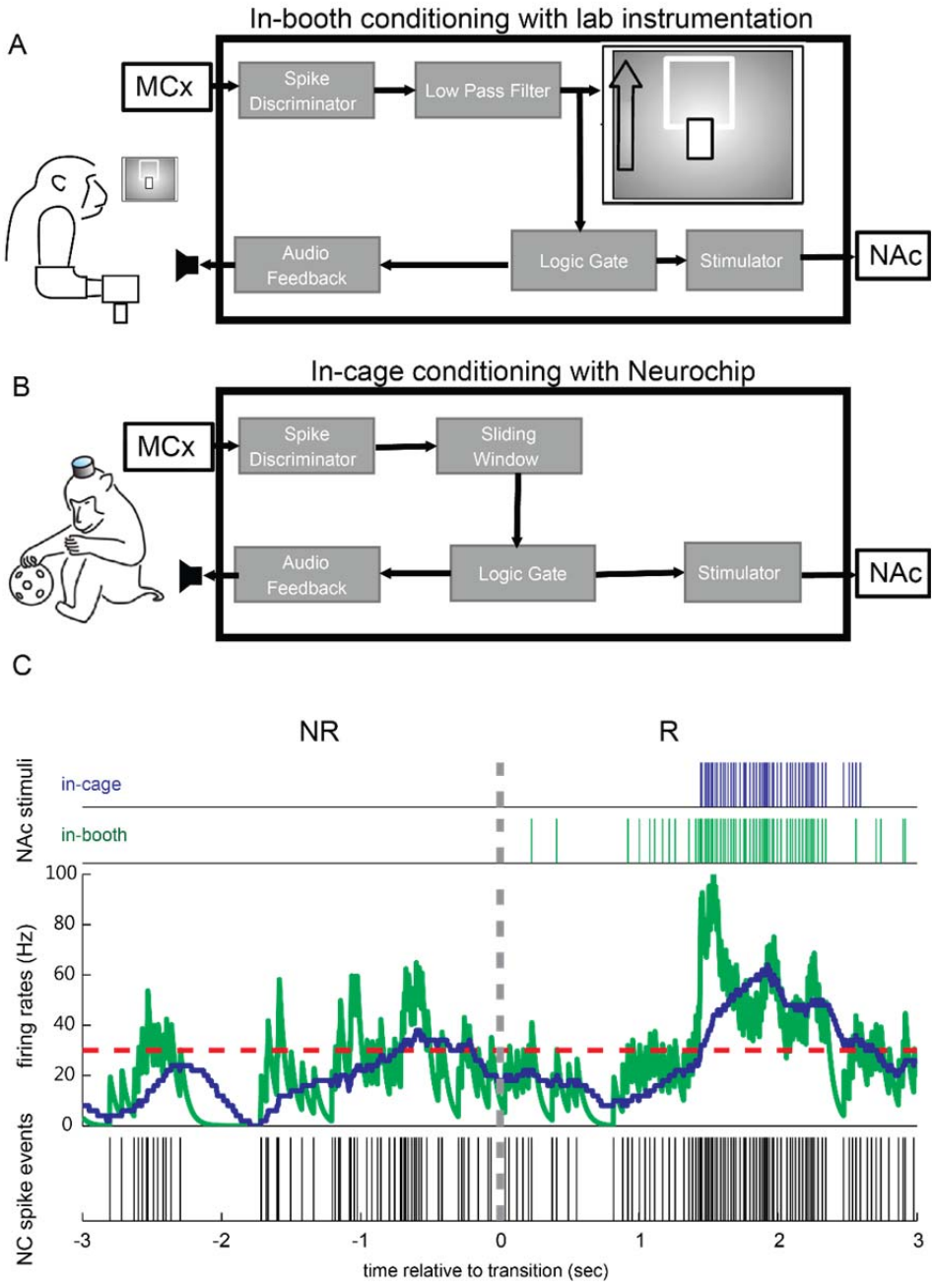


Figure 2. Experimental Conditions. **A.** Schematic of unit conditioning in booth. Activity of motor cortex (MC) cell generated pulses that were low-pass filtered and controlled cursor position on a screen. A logic gate triggered pulses when firing rate exceeded a threshold (green traces in C). Pulses stimulated nucleus accumbens (NAc) and auditory feedback tones. **B.** Schematic of unit conditioning during free behavior in cage. The Neurochip was programmed to detect spikes and compile a running average of rate; when this exceeded threshold, pulses triggered stimuli to NAc (blue traces in C) and auditory clicks. **C.** Conversion of NC spike events (bottom) to NAc stimuli (top) as firing rates exceeded threshold (red dashes) for in-booth and in-cage conditioning (green and blue, respectively)

During alternating R/NR conditioning, we approximated instantaneous firing rate in real-time using two methods, depending on the environment (Figure 2). For most in-booth sessions, spikes were discriminated with two time-amplitude windows and each spike event triggered a 1 ms wide square pulse. The pulse train output (Fig. 2C bottom) was low-pass filtered (τ = 50 ms) and amplified using an analog leaky integrator. These operations produced a continuous signal (Figure 2C, green trace) that controlled cursor movements on the display in front of the animal, providing visual feedback of rate relative to target (Figure 2A). When the activity-controlled cursor entered the target, all subsequent in-target spike events triggered stimulation of the reinforcement site. Stimulation events were often also used to trigger auditory clicks. We initially set the target position just above baseline firing rate, and gradually raised its position over the course of conditioning to elicit higher spike rates. Targets were presented only during R periods of the alternating R/NR task.

For in-cage sessions (Figure 2B), we pre-programmed the NC to perform a real-time sliding window operation to estimate instantaneous spike rate (Figure 2.C, blue trace). The NC counted the number of spike events within a 500 ms wide moving window that advanced every 10 ms. The NC delivered spike-triggered stimuli on spike events that occurred when this estimated rate exceeded a threshold frequency (Figure 2C, red dashed line). Threshold was determined from force target-tracking or in-booth R/NR task response averages that revealed baseline and maximum firing rates of the particular cell. Typically, in-cage stimulation thresholds were set at 75% of the observed maximum firing rate of the candidate cell. In later sessions, the NC governed operant conditioning sessions both in the training booth and cage, to directly compare the effects of environment.

Prior to conditioning, durations of alternating R and NR periods were randomly selected, with replacement, from uniform distributions spanning 1 to 2 min for R and 3 to 5 min for NR. We employed random period durations, within limits, to reduce the monkeys' ability to anticipate transitions in the reinforcement schedule.

Data Analyses

Time series analysis detects rate changes in the alternating R/NR task

To determine whether firing rates during R and NR periods were significantly different, we calculated time-averaged rates during R and NR periods over each conditioning session (e.g., figure 5, left) and pooled them to show rate difference between R and NR periods overall (figure 5, right). Confidence intervals for the time averaged means were computed using a non-parametric bootstrap method based on the Poissonian property of independent inter-spike intervals (ISIs) (Dayan and Abbott, 2001). Specifically, ISIs from each period were randomly drawn with replacement and then summed until their cumulative duration nearly matched the period duration. The number of events comprising the drawn sample divided by period duration produced an estimate of time-averaged rate. Repeating the process 499 times generated a bootstrap distribution of time-averaged rates from which the surrounding 95% confidence interval was determined for each period (T-bars, figure 5). To detect statistically significant patterns in neural activation produced by reinforcement, we computed serial correlation and von Neumann ratio test statistics on the sequence of alternating R-NR-R… time-averaged rates for each conditioning session. These statistics and methods of significance appraisal are described in detail in (Eaton, 2014).

Peri-transition spike activity plots and spike shuffling

To document changes in neural activity around the transitions between R and NR periods we compiled peri-transition histograms of spike activity (figures 6-8). Snippets of the spike trains from 75 seconds before to 75 seconds after each transition were extracted and combined into peri-event spike histograms (binwidth = 50 ms) (e.g., figure 6, black histograms) and consolidated into a single dense train that was convolved with a Gaussian kernel (figure 6, solid red). Shuffled spike rates were obtained by drawing samples with replacement from the list of observed spike events and similarly smoothed (figure 6, solid gray). The process was repeated 199 times to generate a bootstrapped distribution of rate traces from which confidence interval boundaries were calculated (Davison and Hinkley, 1997; Eaton, 2014) (figure 6, dashed gray). Domains in which the

observed rates diverged outside the confidence interval of the shuffled rates indicate features in peri-transition spike activity that could not be explained as random fluctuation.

Results

Accumbens stimulation reinforces target-tracking behavior

We tested the efficacy of candidate reinforcing sites by measuring the monkeys' rate of responding in a manual force target-tracking (FTT) task which they had been trained to perform with applesauce reward. At effective sites trains of stimuli (25 1-mA-pulses at 50 Hz) delivered upon completion of 1-second force holds reinforced further responding. As shown in Figure 3A, the monkey responded at regular rates during reinforcement (R) periods when target completions triggered trains of brain stimulation reward (BSR). Response rates during R periods were significantly higher ($p < 0.001$) compared to interleaved periods during which only feedback was presented and no stimulation was delivered (FO-periods). At the onset of the R periods, which were cued by a tone, response rates often returned quickly to those of the previous R-period. As a comparison, FTT task response rates for applesauce reward typically ranged between 10 and 13 responses per minute for the three monkeys.

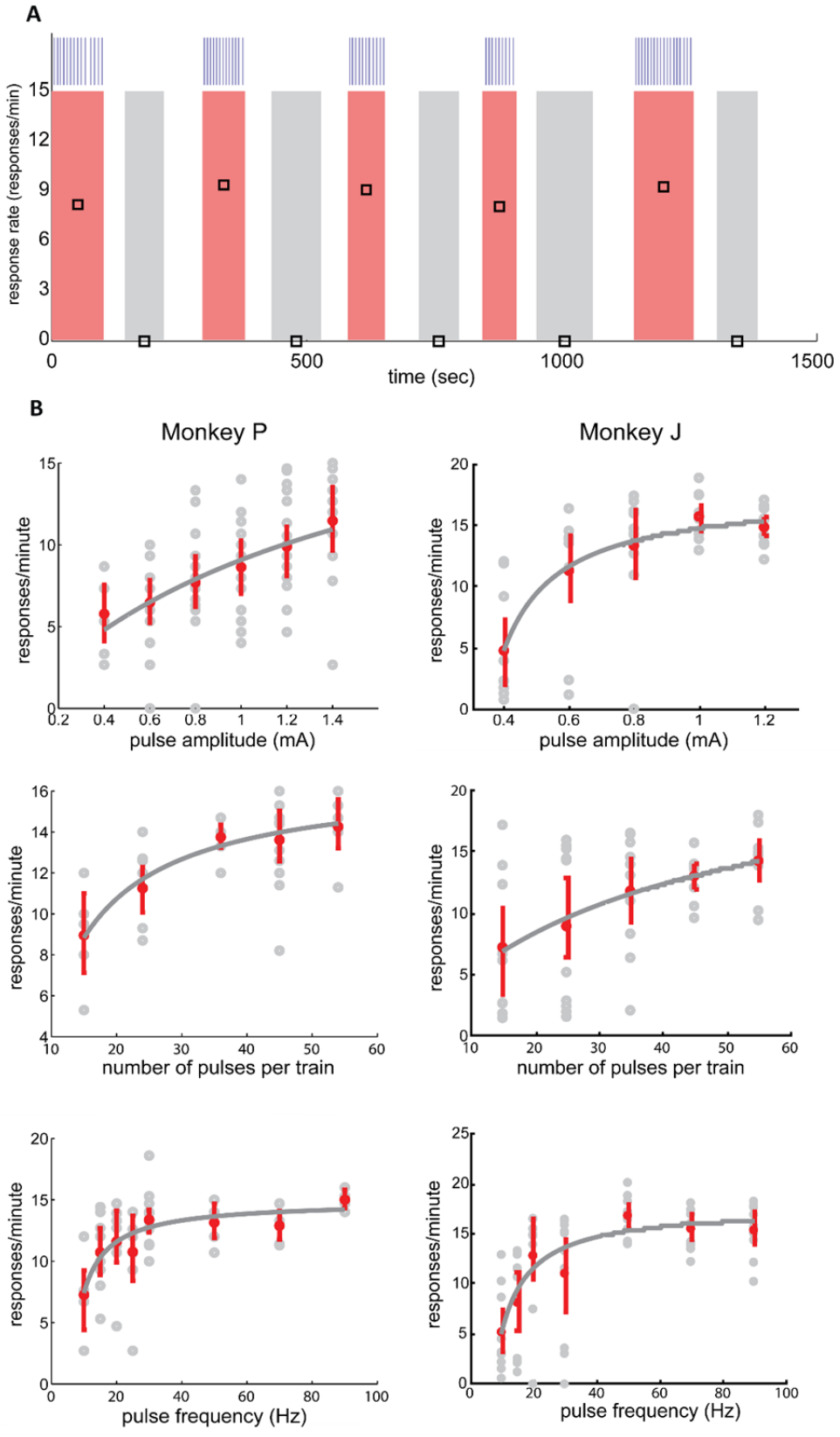


Figure 3. NAc stimulation reinforces target-tracking behavior. **A**. Average response rates (black squares) during

reinforced (R, pink) and feedback only (FO, gray) periods of the wrist force target-tracking task. Clear intervals are non-reinforced periods. Blue ticks (top) mark trial completions. During R periods, each completed flexion or extension hold triggered behaviorally-reinforcing brain stimulation (BSR).
B. Rate of target-tracking responding increases monotonically as a function of increasing BSR intensity**.** Data from monkey P (left column) and monkey D (right). Force target-tracking response rates (gray circles) and response rate means and standard errors (red) are plotted as functions of one varied stimulation parameter: current intensity (top), number of pulses (middle) or frequency (bottom). In each case the other two parameters remained fixed, at 1 mA, 50 Hz or 25 pulses per train. Gray curves depict response rates predicted by regression-fitted Law of Effect model using parameters in Table 1.

Target-tracking rates as function of BSR parameters

In order to determine appropriate stimulation parameters for conditioning cortical cell activity, we documented rates of target-tracking responses for different values of three BSR parameters: current intensity, pulse frequency and number of pulses per stimulus train. Each of these parameters was varied while the other two remained fixed. Fixed values were 1 mA for current intensity, 50 Hz for pulse frequency and 25 pulses per train. For each varied parameter, the values in the desired range were repeated 10 times, delivered in a randomized sequence, to eliminate possible "history effects".

Figure 3B. depicts target-tracking response rates as a function of each stimulus parameter in monkeys P and D. In all cases, the response rates R as a function of the tested stimulation parameter r were well characterized by nonlinear-regression-fitted curves of the Law of Effect model:

$$R(r) = \frac{k(r-r_{th})}{(r-r_{th})+r_e} \qquad \text{Eq. 1}$$

where r_{th} is the threshold level, or lowest value at which the stimulus parameter supported self-stimulation, and r_e represents the aggregate reinforcement for all non-operant responses (Herrnstein, 1970). Table 1 summarizes fit statistics for each of the plots. The response curves indicate that ~80-90% of maximal responding (horizontal asymptote of each plot) occurred for stimulation parameters 1 mA and 50 Hz.

In subsequent cell and muscle conditioning experiments pulse amplitude was set to 1 mA. Bursts of elevated spike rates triggered pulse trains at frequencies approaching 50 Hz. For slowly firing cells (e.g. <10 Hz), multiple stimulus pulses (delivered at 50 Hz) were triggered for each RCST stimulus event.

264

265 Table 1. Single response Law of Effect model fit parameters and statistics

266

Varied parameter	*Monkey*	*k*	r_e	r_{th}	*MSE*
Pulse amplitude (mA)	P	22.45	1.47	0.00	0.15
	D	17.64	0.13	0.35	0.31
No. pulses per train	P	16.65	7.31	6.75	0.16
	D	23.72	37.00	0.00	0.04
Pulse frequency (Hz)	P	14.89	4.00	6.05	0.35
	D	17.74	6.86	7.32	1.46

267 *Table 1. Parameters fit to single response Law of Effect model (eq. 1.). The fit parameters are: k, the maximal response rate asymptote*
268 *(responses/min.),* r_e*, the aggregate reinforcement for all non-operant responses, and* r_{th}*, the threshold level, or lowest value at which the*
269 *varied stimulus parameter supported self-stimulation. MSE: mean squared error of the model fit using non-linear regression.*

270

271 **Muscle activity reinforced during free behavior with BSR**

272

273 To confirm the efficacy of BSR sites during free behavior we tested in-cage conditioning using EMG activity of
274 upper limb muscles as the operant in Monkey P. The time-amplitude window discriminator detected biphasic
275 waveforms in the multi-unit EMG signal (Figure 4A right) and generated acceptance pulses whose frequency
276 increased with intensity of muscle contraction. During reinforcement periods the mean rates of biceps EMG-
277 generated pulses were significantly larger than during intervening non-reinforced periods (Figure 4A left), and
278 the monkey was observed to flex his arm during reinforcement periods. With biceps conditioning these
279 differences were maintained for up to 20 hours of conditioning. Significant differences were also seen with
280 triceps conditioning (Figure 4A).

281

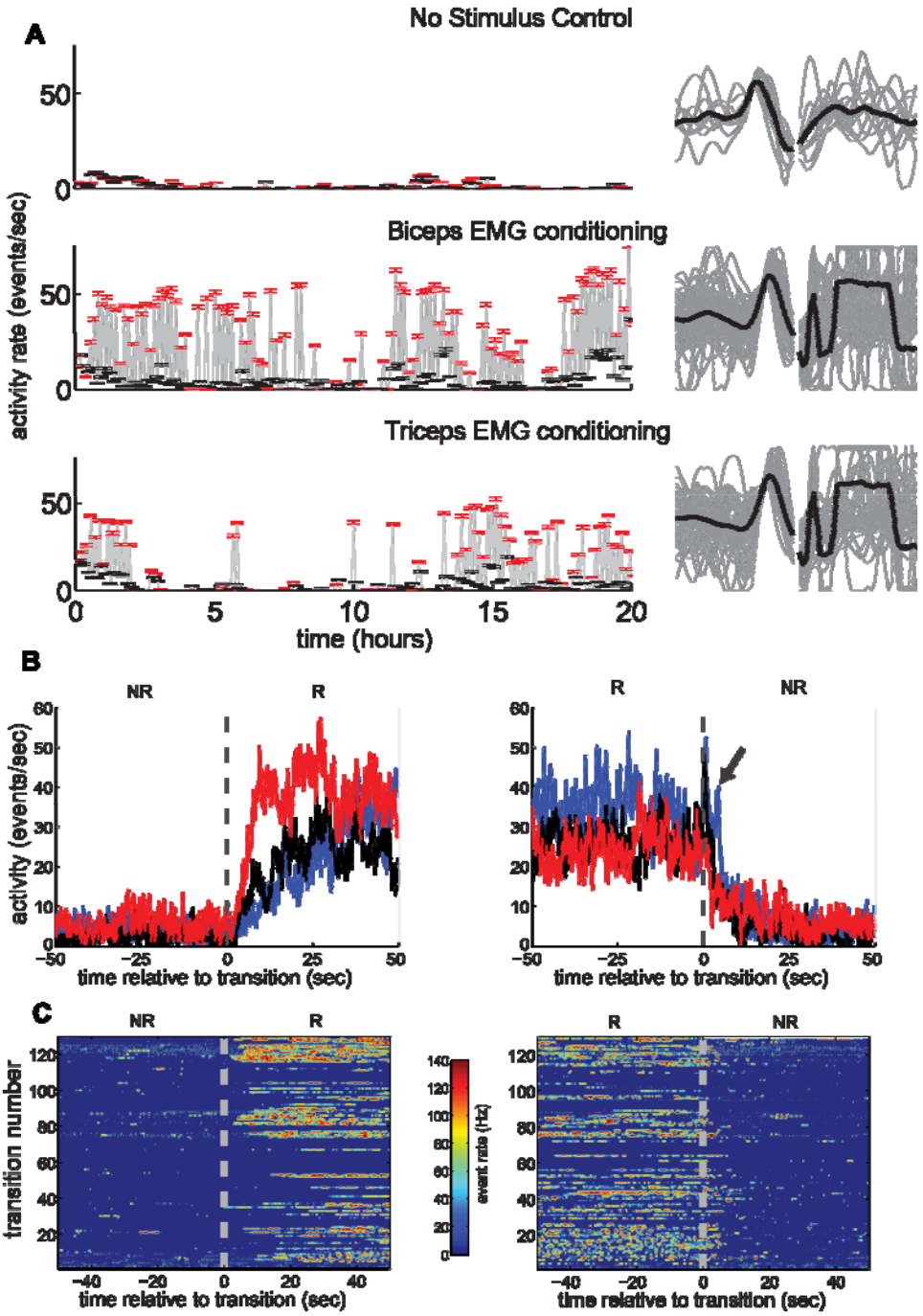


Figure 4. **A.** NAc stimulation reinforces muscle activity in-cage. Top: Baseline control session in which no stimulation was delivered during either "R" (red) or NR (black) periods. Each data-point indicates mean EMG activity over 5 minutes, and surrounding whiskers mark standard error boundaries. Middle and Bottom: EMG-contingent stimulus pulses delivered to NAc during R- periods. Right: EMG potentials that triggered stimulation over each 20-hour session (gray), and their averages (black). Accepted biphasic patterns are followed by artifacts from triggered stimuli. **B.** Averages of biceps muscle activity surrounding NR-R schedule transitions (left) and R-NR transitions (right), shown separately for first third (blue) middle third (black) and last third (red) of the session. **C.** Peri-transition biceps activity during the 20-hour conditioning session, showing NR-R (left) and R-NR (right) transitions. Ordinates count the transitions over the course of the 20-hour session. The color indicates the

rate of biceps EMG activity (see scale).

The transitions between periods of reinforcement (R) and non-reinforcement (NR) showed further evidence of learning to perform the biceps responses. Separate averages around these transitions for the initial, middle and final third of the session (Figure 4B) show progressive changes in responding over the course of the conditioning session. For the NR-to-R transitions rate increases were comparatively low and gradual during the first 6 hours, moderate during the middle period, and greatest and fastest during the last 6 hours. Interestingly, the R-to-NR transitions exhibited a brief increase in responding after the cessation of reinforcement for the first and middle thirds of conditioning (arrow), and no such peak in the last third. Since the monkey had no discriminative stimulus to distinguish R and NR, this behavior is consistent with initial attempts to sustain reinforcement that drop out after sufficient experience with the transition. The raster plots in Figure 4C show color-coded rates for the individual transitions and their variability in more detail. These data confirm that BSR can effectively reinforce an operant, muscle activity, for long periods of time during free behavior.

Overview of cell conditioning sessions

Table 2 summarizes results from all sessions in which cortical cell activity was conditioned with BSR for the three monkeys, categorized by environment: booth or cage. Given sufficient stability and unit isolation, we often conditioned the same cell over repeated sessions. Determining the appropriate conditioning procedures included about 70% of in-cage attempts that were deemed invalid for one or more of the following reasons: 1) Neurochip malfunction, 2) loss of action potential isolation, and 3) improper conditioning parameters.

Table 2. Summary of all effects from cortical spike-triggered BSR conditioning attempts.

	in-booth				in-cage			
	effect type			% success	effect type			% success
	+	0	-		+	0	-	

Monkey P	cells:	7	2	1	70.0	--	--	--	--
	sessions:	7	4	4	46.7	--	--	--	--
Monkey D	cells:	21	3	2	80.8	3	1	0	75.0
	sessions:	38	12	4	70.4	3	2	0	60.0
Monkey J	cells:	20	2	0	90.9	6	2	1	66.7
	sessions:	42	14	1	73.7	18	9	1	64.3
Total	cells:	48	7	3	82.8	9	3	1	69.2
	sessions:	87	30	9	69.0	21	11	1	63.6

Table 2. Summary of all effects from cortical spike-triggered BSR conditioning attempts across subjects, cells, sessions and conditioning environments. Key: "+"denotes statistically significant increases during R-periods compared to NR-periods, "0"s mark cases where no significant changes were observed between R and NR activities and "–" indicate cases when NR-period spike rates were significantly greater than R-period rates. Often the same cell underwent conditioning in multiple sessions. In this table, a given cell was tallied as generating a positive effect though it may have produced null or negative effects in other conditioning attempts. Cells that showed null (0) and negative (–) effects over multiple sessions were categorized as null.

Spike-triggered NAc stimulation reinforces increased motor cortex cell activity

During reinforcement periods the monkeys received spike-triggered BSR when the instantaneous spike rate exceeded a pre-determined threshold. Table 3 summarizes conditioning parameters used for each of the illustrated sessions.

Table 3. Summary of conditioning parameters used for each example conditioning session.

Session Key	Date	Environ-ment	Conditioning equipment	R-period duration (min)	NR-period duration (min)	Rate threshold (Hz)	Trigger-to-stim. pulse Ratio	Sensory Feedback	Figures
P:1-Biceps	11/24/2007	cage	NC1	5	5	none	1:1	none	4
P:2-Triceps	11/25/2007	cage	NC1	5	5	none	1:1	none	4
D:1-1	5/24/2010	booth	rack	1-2	3-5	20	1:1	VC, AC1	5A,9
D:2-2	9/3/2010	booth	rack	1-2	3-5	15	1:3	VC, AC1	5B,6A,9
J:1-1	11/10/2012	booth	rack	1-2	3-5	30	1:1	VC, AC1	5C,6B,9
J:2-1	11/28/2012	cage	NC2-HV	1	5	50	1:1	AC2	5D,6C,7,9
J:3-1	2/22/2013	booth	NC2-HV	2	5	60	1:1	AC2	8
J:3-1	2/22/2013	cage	NC2-HV	2	5	60	1:1	AC2	8

Table 3. Summary of conditioning parameters used for each illustrated conditioning session. In the table above, we list session keys (in the format: Monkey:Session-Cell), dates, conditioning environments, equipment used, schedule durations, rate thresholds , trigger-to-stimulus pulse ratios, and sensory feedback used and text figures for each of the example sessions depicted in the results. In all sessions, BSR stimuli were delivered in 0.2 ms wide, biphasic square pulses at 1 mA intensity. The conditioning equipment used are the NC1: Neurochip, Rack:, rack-mounted instrumentation, and NC2-HV: the High Voltage Neurochip2.The feedback types are: VC: visual cues in the form of a rate-guided computer cursor, AC1: auditory cues in the form of spike-triggered clicks produced by rack-mounted equipment, AC2: clicks on BSR pulse delivery generated by the Neurochip2-HV.

Figure 5A-C shows average motor cortex neuron spike rates during three representative conditioning sessions performed in the training booth. Robust increases in firing rates were observed during R-periods as compared to the intervening NR-periods, showing successful acquisition of the neural operant. In all plots, rates were significantly greater in R than NR periods, as indicated by predominantly non-overlapping confidence intervals. Figure 5D shows an in-cage conditioning session in which monkey J moved freely about his home cage and the Neurochip2 delivered RCST accumbens stimulation in an alternating R/NR schedule over 8 hours. Average firing rates were statistically greater in R-period compared to NR periods; however these differences were smaller than those observed for typical in-booth-conditioning sessions.

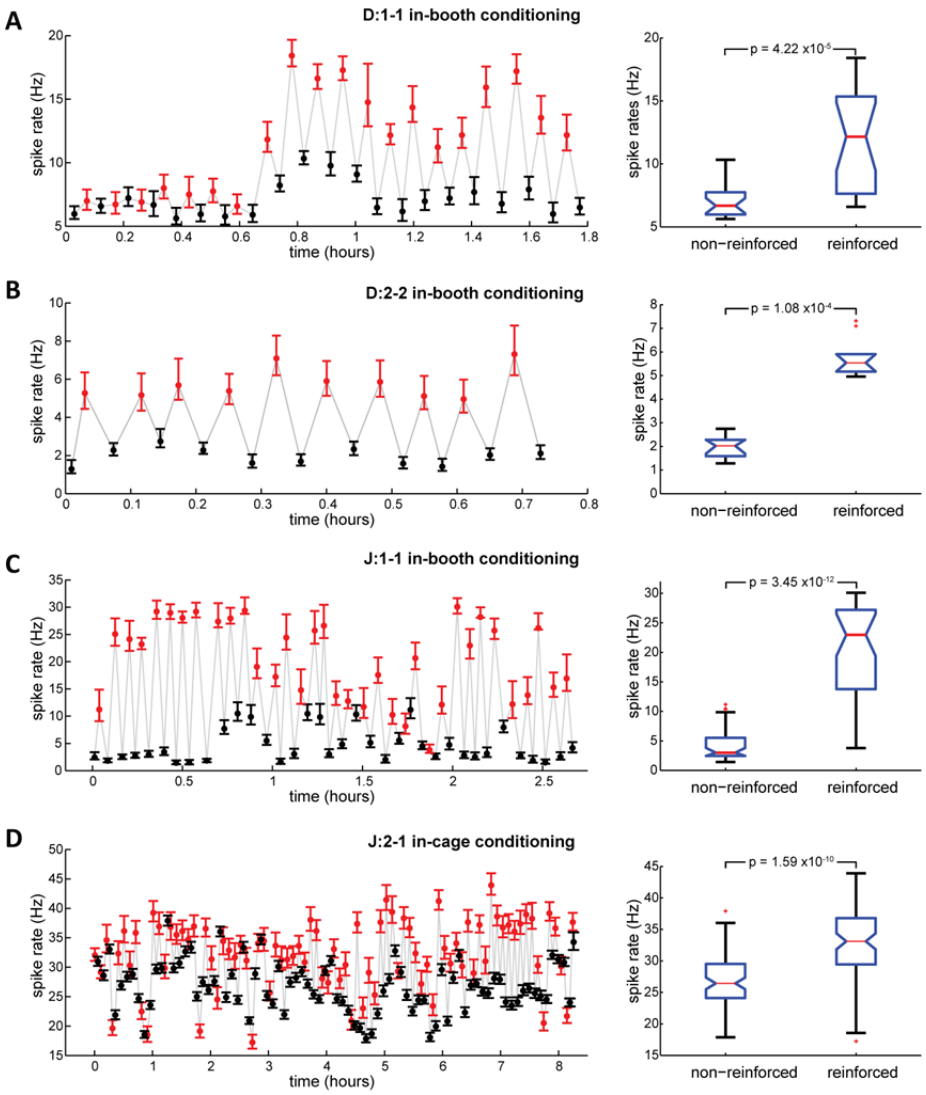


Figure 5. Response rates for cortical neurons during in-booth conditioning (A-C) and in-cage conditioning (D). A and B. initial and later sessions with monkey D. C. in-booth session with monkey J. D. in-cage session during free behavior with monkey J. Left: Points mark average rates during reinforced periods (red) when RCST stimulation was available and during non-reinforced periods (black). Bars denote 95% confidence intervals. Right. Statistics of cortical cell spike rate during R and NR periods shown as box plots. In each box, the central red line marks the distribution median, and blue box extremities depict upper (75%) and lower (25%) quartiles. Red crosses plot rates with values outside of whisker boundaries. Notch height shows approximate limits of confidence intervals about their median at the 5% significance level.

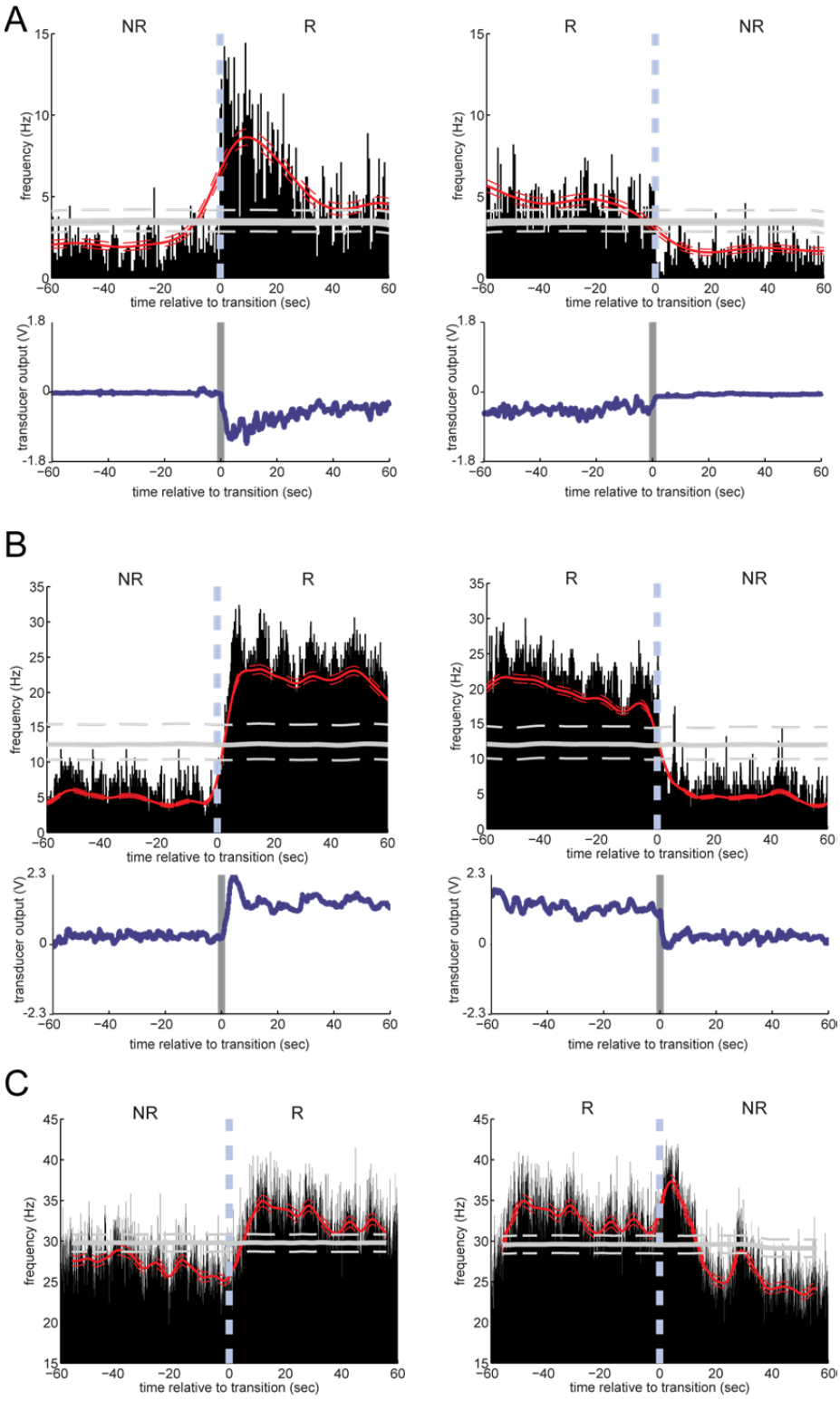
The alternating rate patterns described above give rise to robust, statistically significant time series measures, namely serial correlation and von Neumann's ratio (Eaton, 2014). Alternating rates are obvious from inspection of in-booth conditioning sessions, but are less apparent for the in-cage session. Serial correlation and

von Neumann's ratios measure pattern in time series from which statistical significance can be approximated through randomization and Monte Carlo approximation methods. These analysis techniques confirm significant patterns in these series of time-averages that might otherwise not be evident (Eaton, 2014).

Box plots (Figure 5, right) summarize distributions of reinforced and non-reinforced time averages across each session. For both monkeys, NR distributions have lower medians and were less variable than the R group distributions. These differences are statistically significant in all four examples as assessed by the Kruskal-Wallis test.

Peri-transition activity patterns

For further insight into behavioral mechanisms we documented the changes in firing rates associated with transitions between R and NR periods. Figure 6 shows histograms and smoothed rate traces of neuron spike trains during NR-R and R-NR transitions. For comparison, the overall average rates and 95% confidence intervals are illustrated by grey solid and dashed lines, respectively. Statistically significant deviations from chance occur where the red rate trace exceeds the "chance band". Two sets of peri-transition averages, one for monkey J and one for monkey D exemplify robust rate increases observed across NR-R transitions while the animals underwent RCST stimulation conditioning while under restraint in the training booth. In session J1-1 (Fig. 6B), monkey J produced a four-fold increase in motor cortex cell spike rate and kept rates elevated, on average, for the full duration of reinforcement. During in-booth sessions, activity peaked early, usually within 10 seconds following the NR-R schedule transition, and then decayed over the remainder of each reinforced period. During in-cage conditioning activity peaked later in the R period. Spike activity dropped quickly following R-NR transitions both in-booth and in-cage. However, as shown in Figure 6C, NR spike activity tended to be more variable in the cage than in the booth.



380

381 Figure 6. Peri-transition histograms of neuron spike activity during in-booth (A,B) and in-cage (C) conditioning.
382 For each session, spikes occurring during two-minute intervals straddling NR-R transitions (left) and R-NR
383 transitions (right) are pooled and binned into histograms above. Light-blue dashed vertical lines at t = 0 mark

onset and offset of activity-dependent BSR. The point-density estimate of spike rate (thick red line) and its 95% confidence band limits (dashed red lines) overlay corresponding histograms. Horizontal gray lines show averages (solid) of sweeps, and surrounding 95% confidence intervals (dashed) after spike shuffling. Red line shows smoothed firing rates where the smoothing kernel width was determined by overall spike rate (Davison and Hinkley, 1997). The slow early rise in A is a result of low NR baseline rates followed by an abrupt increase. Bottom: averages of flexion-extension torques recorded concurrently. Depicted here are peri-transition activities from sessions: A. D2-2 and B. J:1-1. C. Peri-transition histograms of spike activity conditioned in-cage during session J:2-1. The NC2 delivered rate-contingent, spike-triggered BSR in the alternating R/NR schedule.

Instrumentation in the training booth allowed us to record wrist torque during unit conditioning. In all examples, motor cortex neurons modulated their activity during dynamic and/or static phases of the force target tracking task. Peri-transition averages of the isometric torque signals show increased torques during R periods that accompanied spike rate increases and corresponding reduction of torque generation during NR with lower cortical spike rates (Figure 6A,B).

Consistent with the parallel analysis of sequential time-averages (Figure 5, left), the increases in spike rates across NR-R transitions were greater in cells conditioned in-booth than in cells conditioned in-cage.

Rate changes of motor cortex cell spike activity conditioned in-cage

As with EMG activity (Figure 4), for in-cage unit-conditioning the relative increases in BSR-reinforced spike activity were smallest, compared to NR-period activity, during the first third and greatest during the final third of the session (Figure 7). A transient increase in spike rate also followed R-NR transitions, when high-frequency spike bursts no longer triggered NAc stimulation. A similar post-extinction burst effect was seen in R-NR peri-transition averages of in-cage conditioned biceps activity (Figure 4B) of the first and middle third session averages. Unlike muscle conditioning however, the extinction burst in spike activity, though markedly reduced, did not completely disappear during the final third of the unit-conditioning session.

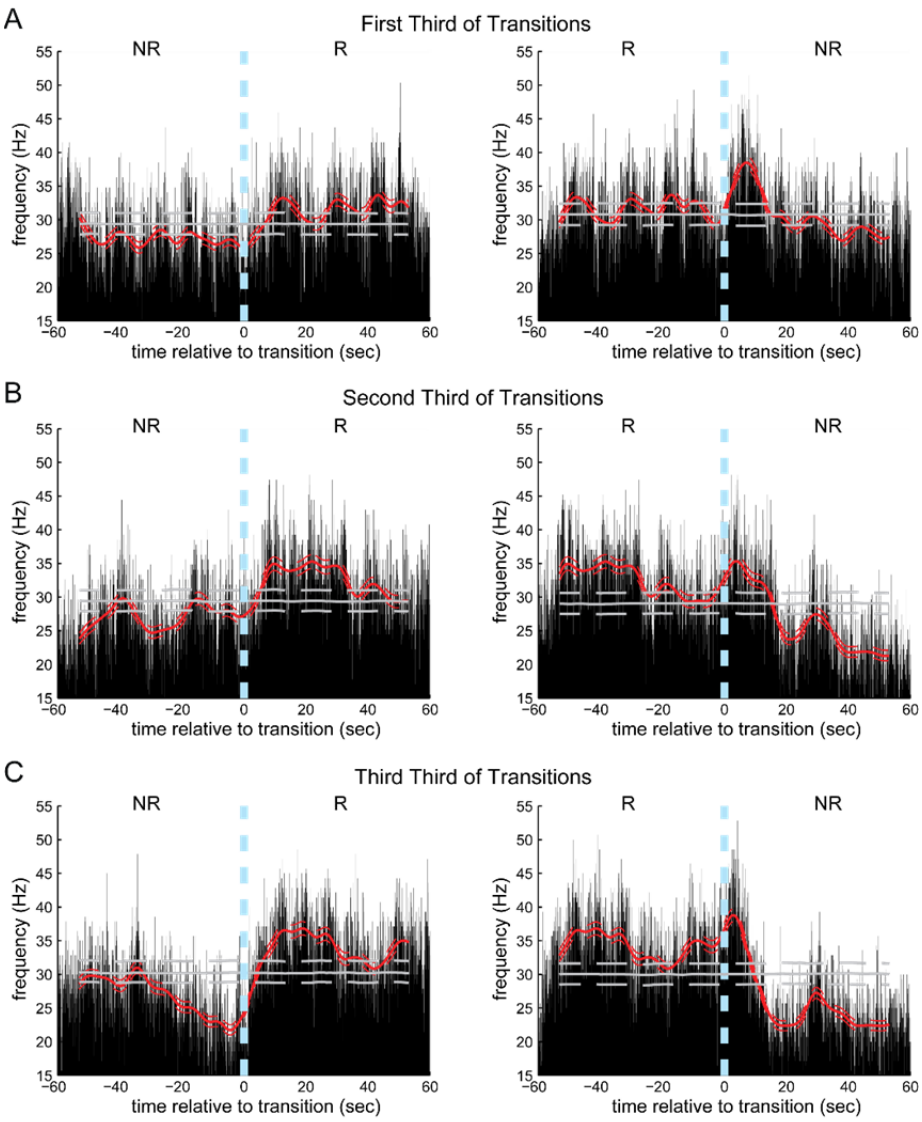


Figure 7. In-cage-conditioned spike activity grouped by first third (A), second third (B) and final third (C) of session. Spike activity from the example shown in Figure 6C.

Cell conditioned in both environments reveals greater efficacy of in-booth conditioning

The above evidence suggests that greater conditioning effects were obtained during in-booth conditioning with restraint and visual feedback than during in-cage sessions with free behavior. This could have been due to the slight difference in reinforcement paradigms (Figure 2) as well as environment. For a definitive comparison, we conditioned the same cell, using identical conditioning parameters, both in the training booth and as monkey J

moved freely about his cage. Figure 8 shows rates when spikes from a motor cortex neuron triggered NAc stimulation during elevated firing rates. Stimulation was available during 2-min R periods alternating with 5-min NR periods. During the first hour, the monkey underwent unit conditioning while he moved freely about his cage; he was then transferred within 6 minutes to his training booth and restrained. The Neurochip delivered identical conditioning stimulation in both environments. During reinforcement periods, single 1mA biphasic pulses were delivered to NAc on each event that exceeded 30 counts within a 500-ms-wide sliding window updated every 10 ms. Figure 8A plots cell spike activity as time-averaged rates. Horizontal dashed lines show group means of reinforced and non-reinforced intervals for each environment (red and black respectively). The Neurochip generated an auditory click on each stimulation pulse event to provide a discriminative stimulus. No visual feedback was provided in either environment.

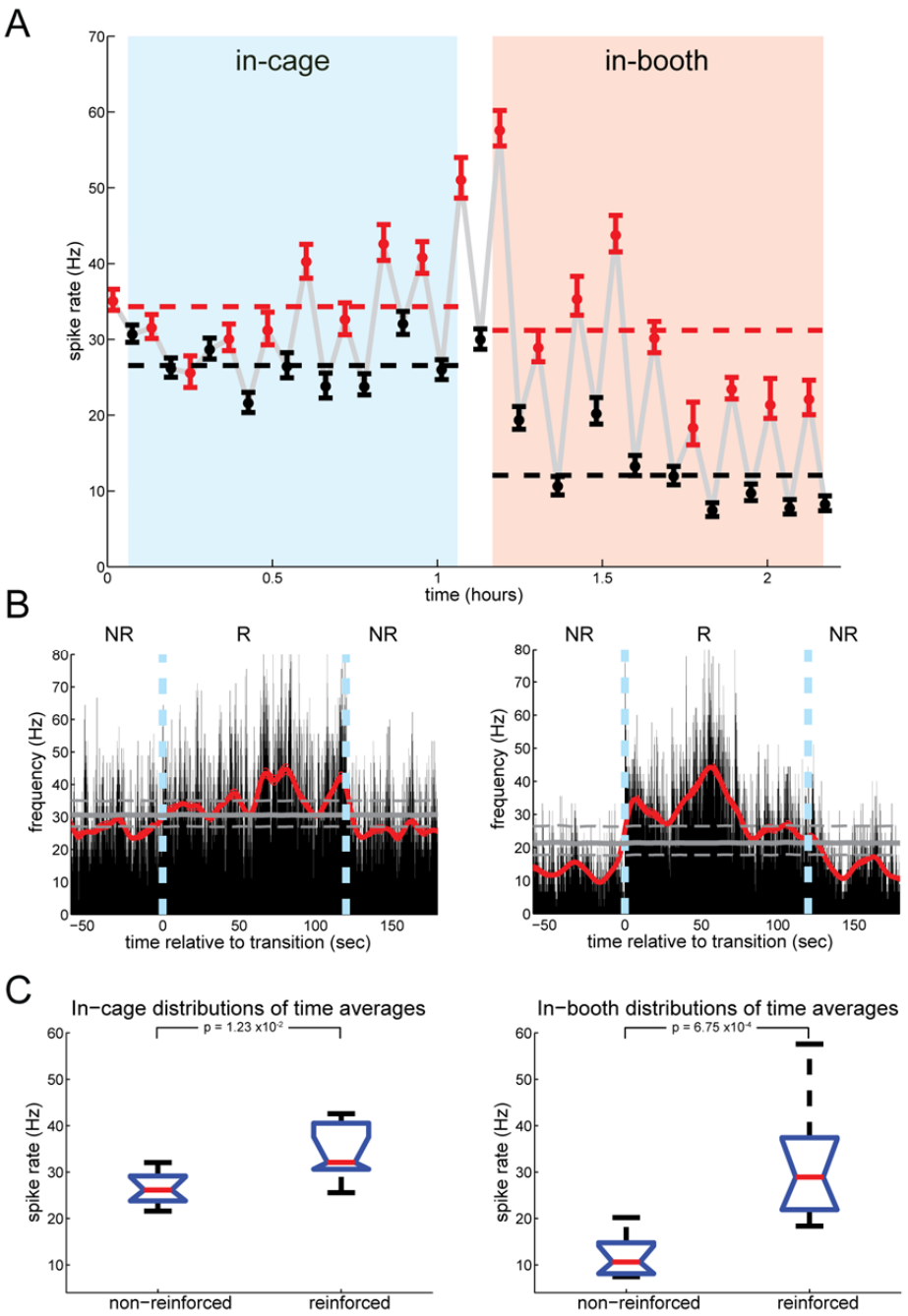


Figure 8. Rates of motor cortex neuron conditioned in-cage and in-booth with identical parameters. A. Average firing rates during R (red) and NR (black) and surrounding 95% confidence intervals. The monkey was in his home cage (light blue background) during the first hour and then quickly transferred to the training booth during second hour (tan). B. Peri-transition spike averages compiled during conditioning periods in-cage (left) and in-booth (right). Horizontal lines show overall mean rate for each environment. The Neurochip controlled conditioning in both environments; it ran continuously during the 6 minute transfer interval, and continued uninterrupted through the entire session. C. Statistics of cortical cell spike rate during R and NR periods shown as box plots for in-cage (left) and in-booth (right). The p values were obtained from the Kruskal-Wallis test.

The progression of alternating time-averages of reinforced and non-reinforced cortical cell firing rates show statistically significant increases during periods of BSR reinforcement compared to the intervening NR periods, both in the training booth and the end of in-cage conditioning. Comparisons between distributions of pooled R and NR time averages show statistically significant increases during reinforcement (Figure 8C), in both the cage and the booth. The group median of NR period averages during in-cage conditioning (25 Hz) was substantially greater than the median of the NR group during in-booth conditioning (13 Hz), indicating higher baseline rates during free behavior. Peri-transition firing rates (Figure 8B) also show higher baseline activity during in-cage than in-booth NR periods and show that cell firing peaked midway through the two-minute reinforcement interval.

Accumbens stimuli do not evoke cortical responses

Recent anatomical investigations (Miyachi et al., 2005, Miyachi et al., 2006) suggest a pathway through which input from the NAc could reach primary motor cortex more directly than the well-established striatal-pallidal-thalamo-cortical circuit (Alexander et al., 1990) (Parent and Hazrati, 1995). To address this possible confound of direct stimulus-evoked effects in cell firing, we delivered continuous 5 Hz test pulses to the BSR site while recording spike activity of the candidate cell prior to each conditioning session. None of the candidate cells exhibited statistically significant increases in firing probability at any latency between 0 and 200 ms following single-pulse stimuli delivered to NAc at the current intensity (1 mA) used for BSR. The four representative cases in Figure 9 show that the 95% confidence intervals surrounding kernel-smoothed traces of the observed spike event sequences (red) did not exceed chance levels (gray), indicating that the modest transient fluctuations in spike probability in these histograms did not achieve statistical significance. Thus, striatal-cortico linkage did not contribute directly to increases in cortical cell spike activity during unit conditioning with BSR.

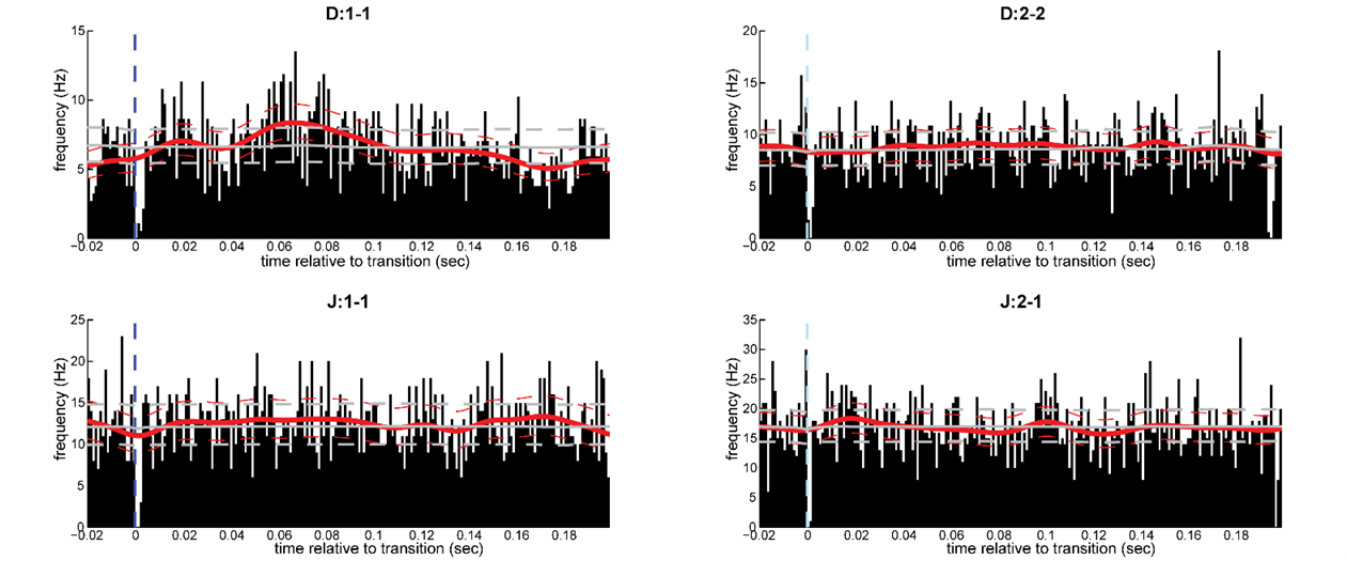


Figure 9. Peri-stimulus spike histograms compiled from spike trains surrounding single-pulse stimulation of NAc reinforcement sites. Histograms were compiled just prior to example experiments depicted in Figure 5A and Figure 5B using the same stimulation intensity. Vertical dashed lines depict stimulus delivery. Histogram bin width: 2ms. Point density average (thick red line) and limits of the 95% confidence band (red dashed lines) depict kernel-smoothed approximates of peri-stimulus mean spike rates.

Discussion

This study shows that firing rates of motor cortical neurons and muscle activity can be operantly reinforced through delivery of rate-contingent stimulation of ventral striatum in non-human primates. We identified BSR sites in NAc whose stimulation reinforced performance of a target-tracking task with reward efficacy comparable to fruit sauce. Systematic testing of stimulus parameters (width, amplitude and frequency) with the FTT task demonstrated response rates consistent with the Law of Effect (Herrnstein, 1970). Our stimulation of NAc probably activated fibers that evoked dopamine release, including fibers from the medial forebrain bundle, which connects the ventral tegmental area to NAc and whose stimulation supports operant responding (German and Fetz, 1976). Axon terminals of the medial forebrain bundle release dopamine within the NAc on receipt of unconditioned rewards (Wise, 1978, Hernandez and Hoebel, 1988). Moreover, the reinforcing effects of stimuli that are normally rewarding, such as food, water, drugs of abuse and stimulation of the medial forebrain bundle,

477 are blocked in animals given dopamine antagonists (Wasserman et al., 1982). A significant proportion of
478 macaque NAc neurons modulated their activity during task-contingent delivery of juice rewards (Apicella et al.,
479 1991). Thus, the reinforcing effects of our stimuli were likely mediated by activating fibers that released
480 dopamine.

481 ## Functional relationships between motor cortex and striatum

482 The functional relations between the ventral striatum and motor cortex have been elucidated by anatomical
483 electrophysiological and behavioral studies. Polysynaptic projections from NAc to motor cortex have been
484 revealed by retrograde transsynaptic transport of rabies virus (Miyachi et al., 2006). Conversely, the motor
485 cortex is one of the cortical areas from which the ventral striatum receives input (Takada et al., 1998), (Tokuno
486 et al., 1999). Simultaneous recordings of cortical surface ECoG and local field potentials in NAc showed evidence
487 for electrophysiological interactions, in a study demonstrating that NAc plays a significant role in recovery of
488 motor function after corticospinal lesions (Sawada et al., 2015). Temporally precise coherence between output-
489 relevant neuronal populations in motor cortex and dorsal striatum developed during learning to control cortical
490 cell activity (Koralek et al., 2013). Despite this evidence for close relations, we found no evidence that our NAc
491 stimuli produced any post-stimulus modulation of motor cortex neurons, indicating that the effect of stimulation
492 on firing rates was mediated by behavioral reinforcement.

493 ## Activity correlated with conditioned neurons

494 While BSR was delivered contingent on increases in firing of a single motor cortex cell, larger neuronal
495 populations would obviously have to be co-activated; in particular, other neurons that provide direct and
496 indirect input to the conditioned neuron would also be recruited to drive its rate increases. Such co-activation of
497 large populations was evidenced by associated muscle contractions and neighboring cell activity. During in-
498 booth sessions the monkey's conditioned changes in neural activity were often correlated with isometric
499 torques produced around the wrist. This is not surprising since the neurons chosen for conditioning were

modulated during the wrist task. A previous study found that chaired animals allowed to move limbs freely generated a variety of movements associated with operant bursts of the same cell (Fetz and Baker, 1973). Given this variability we did not attempt to document the monkeys' movements during the in-cage neural conditioning sessions. A more systematic analysis of movements related to operant bursts during free behavior could be pursued using simultaneous neural and video recordings.

In some sessions the activity of a neighboring cell was recorded simultaneously with the reinforced neuron. As illustrated in Eaton 2014, neurons whose cross-correlograms had central peaks indicative of common synaptic drive from upstream sources to both cells could be coactivated or modulated reciprocally in the R/NR periods. These results are consistent with previous studies of synaptic linkages between motor cortex neurons showing that common inputs are seen for both coactivated and reciprocally activated pairs (Smith and Fetz, 2009).

Comparison of neural conditioning in-booth and in-cage

Learning to control neural activity progressed more slowly during in-cage than in-booth conditioning sessions. In addition, rate increases were smaller and harder to discern for in-cage R-period versus NR-periods. Several differences between the two conditioning environments could have contributed to this disparity. First, during in-booth sessions the monkeys were restrained, with their head and contralateral arms secured. We believe such restraint effectively reduced activity of the movement-related cells during NR periods, providing a lower "baseline" against which increases were measured. Second, most in-booth sessions involved stronger discriminative stimuli (e.g., auditory clicks and a rate-controlled computer cursor) than the barely-audible clicks produced by the Neurochip during in-cage sessions. More intense discriminative stimuli are more likely to be effective secondary reinforcers during the conditioning task. Third, the lack of restraint during in-cage conditioning permitted monkeys to explore a much broader range of motor activities. The greater behavioral repertoire provided more distractions when forming response-reward associations, thus requiring longer time to demonstrate acquisition. In contrast, in the training booth, where monkeys had spent many hours performing the FTT task for both food reward and BSR, monkeys likely drew from a much smaller pool of potential reward-

524 eliciting responses when forming neural-response-reward associations. Fourth, the low-pass filtering of neural
525 activity used for most in-booth experiments may have been more effective than the sliding-window method
526 used for in-cage Neurochip sessions (Figure 2C). This possibility was disproven in a control session in which the
527 sliding-window method was used for both environments: the monkey's performance was still more robust in the
528 booth, where baseline firing rate was lower (Figure 8).
529 Finally, consistent with the r_e parameter of the Law of Effect model, the in-cage environment introduced
530 additional reinforcers – for example, food, toys, presence of neighboring monkeys and grooming activities – that
531 served to increase competing behaviors to the spike-rate operant. As the collective contribution from all non-
532 task reinforcers, r_e, increases, the influence of the task-associated reinforcer, r (BSR in our case), on operant
533 responding is effectively reduced, as shown by the mathematical expression of the Law of Effect for response
534 rate (eq, 1), in which the sum of the two terms $r + r_e$ comprise the denominator. Since fewer non-task-
535 reinforced response alternatives are available to monkeys in the training booth, the Law of Effect predicts that
536 the rewards paired to the operant response should be more effective than in the cage, where there are many
537 distractions.
538 Most of the above reasons that efficacy of conditioning during free behavior would be reduced should also have
539 applied for EMG conditioning. However, increased EMG responses proved quite robust for almost 20 hours
540 (Figure 4A). This difference raises the possibility that conditioning of neural activity might be more difficult than
541 muscle activity; however that conclusion would be contradicted by many successful unit conditioning studies
542 using conventional rewards (Fetz and Baker, 1973, Fetz and Finocchio, 1975, Moritz and Fetz, 2011). It may be
543 possible that task acquisition itself was faster for EMG conditioning specifically in the context of free behavior.
544 Thus, while the target muscles were normally active in the monkey's natural movement repertoire, the relevant
545 neural activity may not have been as readily discoverable in the cage. Since bursts of motor cortex neurons are
546 typically related to many different movements (Fetz and Baker, 1973, Fetz and Finocchio, 1975), these diverse
547 relations could have undermined the acquisition of any particular effective movement. These hypotheses

clearly deserve further investigation.

Investigating neural coding

Reinforcement of neural activity with BSR during free behavior has the potential of investigating mechanisms of neural coding. In contrast to the conventional coding of information in neural firing rates, the hypothesis that information could be coded in the precise timing of spike activity remains to be proven. The operation of such temporal coding would significantly expand the bandwidth for neural computation (Fetz, 1997). While we have demonstrated the ability of BSR to reward increases in firing rates, BSR could also be used to test the volitional control of precise spatiotemporal patterns. If the brain uses such patterns during normal behavior, many of them should be volitionally controllable. The use of BSR to instantly reward the appearance of specific patterns under free conditions would provide ample time for the monkey to discover and repeat the relevant behavioral or cognitive state. This would represent a significant test of the existence of temporal coding in the brain.

Acknowledgements

We thank Steve Perlmutter, Chet Moritz, Timothy Lucas, Andrew Jackson and Yukio Nishimura for surgical assistance. Stavros Zanos helped run the in-cage muscle conditioning experiments. Zachary Roberts and Gerick Lee assisted with monkey handling and recording, and Leah Bakst assisted with analysis. Douglas Bowden and Paul Phillips provided helpful discussion.

Author Contributions

RE and EF conceived the experiments; RE and TL performed experiments and RE, EF and TL wrote the manuscript.

Bibliography

Aflalo TN, Graziano MSA (2006) Partial tuning of motor cortex neurons to final posture in a free-moving paradigm. Proc Natl Acad Sci U S A 103:2909-2914.

Alexander GE, Crutcher MD, DeLong MR (1990) Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. Prog Brain Res 85:119-146.

Apicella P, Ljungberg T, Scarnati E, Schultz W (1991) Responses to reward in monkey dorsal and ventral striatum. Exp Brain Res 85:491-500.

Bichot NP, Heard MT, Desimone R (2011) Stimulation of the nucleus accumbens as behavioral reward in awake behaving monkeys. J Neurosci Meth 199:265-272.

Bowden DM ME, McKeown CW, Baldwin DV (2015) Map of Positive and Negative Reinforcement Sites in the Brain of the Rhesus Macaque (Macaca mulatta). (Center, N. P. R., ed) Seattle, WA.

Briese E, Olds J (1964) Reinforcing brain stimulation and memory in monkeys. Experimental Neurology 10:493-508.

Caminiti R, Johnson PB, Burnod Y, Galli C, Ferraina S (1990) Shift of preferred directions of premotor cortical cells with arm movements performed across the workspace. Exp Brain Res 83:228-232.

Carmena JM, Lebedev MA, Crist RE, O'Doherty JE, Santucci DM, Dimitrov DF, Patil PG, Henriquez CS, Nicolelis MA (2003) Learning to control a brain-machine interface for reaching and grasping by primates. PLoS Biol 1:E42.

Davison AC, Hinkley DV (1997) Bootstrap methods and their application. Cambridge, UK ; New York, NY, USA: Cambridge University Press.

Dayan P, Abbott LF (2001) Theoretical neuroscience : computational and mathematical modeling of neural systems. Cambridge, Mass.: Massachusetts Institute of Technology Press.

Delgado JM, Johnston VS, Wallace JD, Bradley RJ (1970) Operant conditioning of amygdala spindling in the free chimpanzee. Brain Res 22:347-362.

Eaton RW (2014) Operant conditioning of cortical cell and muscle response patters. In: Physiology and Biophysics, vol. Doctor of Philosophy Seattle, WA: University Of Washington.

Fetz EE (1997) Temporal coding in neural populations? Science 278:1901-1902.

Fetz EE (2007) Volitional control of neural activity: implications for brain-computer interfaces. J Physiol 579:571-579.

Fetz EE, Baker MA (1973) Operantly conditioned patterns on precentral unit activity and correlated responses in adjacent cells and contralateral muscles. J Neurophysiol 36:179-204.

Fetz EE, Finocchio DV (1975) Correlations between activity of motor cortex cells and arm muscles during operantly conditioned response patterns. Exp Brain Res 23:217-240.

German DC, Fetz EE (1976) Responses of primate locus coeruleus and subcoeruleus neurons to stimulation at reinforcing brain sites and to natural reinforcers. Brain Res 109:497-514.

Hernandez L, Hoebel BG (1988) Food reward and cocaine increase extracellular dopamine in the nucleus accumbens as measured by microdialysis. Life Sci 42:1705-1712.

Herrnstein RJ (1970) On the law of effect. J Exp Anal Behav 13:243-266.

Hiatt DE (1972) Investigations of operant conditioning of single unit activity in the rat brain. California Institute of Technology.

Jackson A, Fetz EE (2007) Compact movable microwire array for long-term chronic unit recording in cerebral cortex of primates. J Neurophysiol 98:3109-3118.

Jackson A, Mavoori J, Fetz EE (2006) Long-term motor cortex plasticity induced by an electronic neural implant. Nature 444:56-60.

Jackson A, Mavoori J, Fetz EE (2007) Correlations between the same motor cortex cells and arm muscles during a trained task, free behavior, and natural sleep in the macaque monkey. J Neurophysiol 97:360-374.

Koralek AC, Costa RM, Carmena JM (2013) Temporally precise cell-specific coherence develops in corticostriatal

networks during learning. Neuron 79:865-872.
Koralek AC, Jin X, Long JD, 2nd, Costa RM, Carmena JM (2012) Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature 483:331-335.
Mavoori J, Jackson A, Diorio C, Fetz E (2005) An autonomous implantable computer for neural recording and stimulation in unrestrained primates. J Neurosci Methods 148:71-77.
Miyachi S, Lu X, Imanishi M, Sawada K, Nambu A, Takada M (2006) Somatotopically arranged inputs from putamen and subthalamic nucleus to primary motor cortex. Neurosci Res 56:300-308.
Miyachi S, Lu XF, Inoue S, Iwasaki T, Koike S, Nambu A, Takada M (2005) Organization of multisynaptic inputs from prefrontal cortex to primary motor cortex as revealed by retrograde transneuronal transport of rabies virus. J Neurosci 25:2547-2556.
Moritz CT, Fetz EE (2011) Volitional control of single cortical neurons in a brain-machine interface. J Neural Eng 8:025017.
National Primate Research Center UoW (1991-present) **BrainInfo**. http://www.braininfo.org.
Olds J (1958) Satiation effects in self-stimulation of the brain. J Comp Physiol Psychol 51:675-678.
Olds J, Milner P (1954) Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. J Comp Physiol Psychol 47:419-427.
Parent A, Hazrati LN (1995) Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. Brain Res Brain Res Rev 20:91-127.
Rathelot JA, Strick PL (2009) Subdivisions of primary motor cortex based on cortico-motoneuronal cells. Proc Natl Acad Sci U S A 106:918-923.
Rolls ET, Burton MJ, Mora F (1980) Neurophysiological analysis of brain-stimulation reward in the monkey. Brain Res 194:339-357.
Routtenberg A, Gardner EL, Huang YH (1971) Self-stimulation pathways in the monkey, Macaca mulatta. Exp Neurol 33:213-224.
Sawada M, Kato K, Kunieda T, Mikuni N, Miyamoto S, Onoe H, Isa T, Nishimura Y (2015) Function of the nucleus accumbens in motor control during recovery after spinal cord injury. Science 350:98-101.
Smith WS, Fetz EE (2009) Synaptic Linkages Between Corticomotoneuronal Cells Affecting Forelimb Muscles in Behaving Primates. J Neurophysiol 102:1040-1048.
Takada M, Tokuno H, Nambu A, Inase M (1998) Corticostriatal input zones from the supplementary motor area overlap those from the contra- rather than ipsilateral primary motor cortex. Brain Res 791:335-340.
Taylor DM, Tillery SI, Schwartz AB (2002) Direct cortical control of 3D neuroprosthetic devices. Science 296:1829-1832.
Tokuno H, Inase M, Nambu A, Akazawa T, Miyachi S, Takada M (1999) Corticostriatal projections from distal and proximal forelimb representations of the monkey primary motor cortex. Neurosci Lett 269:33-36.
Wasserman EM, Gomita Y, Gallistel CR (1982) Pimozide blocks reinforcement but not priming from MFB stimulation in the rat. Pharmacol Biochem Behav 17:783-787.
Widge AS, Moritz CT (2014) Pre-frontal control of closed-loop limbic neurostimulation by rodents using a brain-computer interface. J Neural Eng 11:024001.
Wise RA (1978) Catecholamine theories of reward: a critical review. Brain Res 152:215-247.
Zanos S, Richardson AG, Shupe L, Miles FP, Fetz EE (2011) The Neurochip-2: an autonomous head-fixed computer for recording and stimulating in freely behaving monkeys. IEEE Trans Neural Syst Rehabil Eng 19:427-435.

Figure Captions:

Figure 1: Co-registration of cranial X-ray, MRI and brain atlas images. A coronal map of the M. Nemestrina brain

was morphed and superimposed on an MRI of Monkey P's brain. Both are positioned over an X-ray image showing the chamber and housed hardware. NAc shown in red.

Figure 2. Experimental Conditions. **A.** Schematic of unit conditioning in booth. Activity of motor cortex (MC) cell generated pulses that were low-pass filtered and controlled cursor position on a screen. A logic gate triggered pulses when firing rate exceeded a threshold (green traces in C). Pulses stimulated nucleus accumbens (NAc) and auditory feedback tones. **B.** Schematic of unit conditioning during free behavior in cage. The Neurochip was programmed to detect spikes and compile a running average of rate; when this exceeded threshold, pulses triggered stimuli to NAc (blue traces in C) and auditory clicks. **C.** Conversion of NC spike events (bottom) to NAc stimuli (top) as firing rates exceeded threshold (red dashes) for in-booth and in-cage conditioning (green and blue, respectively)

Figure 3. NAc stimulation reinforces target-tracking behavior. **A**. Average response rates (black squares) during reinforced (R, pink) and feedback only (FO, gray) periods of the wrist force target-tracking task. Clear intervals are non-reinforced periods. Blue ticks (top) mark trial completions. During R periods, each completed flexion or extension hold triggered behaviorally-reinforcing brain stimulation (BSR).
B. Rate of target-tracking responding increases monotonically as a function of increasing BSR intensity**.** Data from monkey P (left column) and monkey D (right). Force target-tracking response rates (gray circles) and response rate means and standard errors (red) are plotted as functions of one varied stimulation parameter: current intensity (top), number of pulses (middle) or frequency (bottom). In each case the other two parameters remained fixed, at 1 mA, 50 Hz or 25 pulses per train. Gray curves depict response rates predicted by regression-fitted Law of Effect model using parameters in Table 1.

Figure 4. **A.** NAc stimulation reinforces muscle activity in-cage. Top: Baseline control session in which no stimulation was delivered during either "R" (red) or NR (black) periods. Each data-point indicates mean EMG activity over 5 minutes, and surrounding whiskers mark standard error boundaries. Middle and Bottom: EMG-contingent stimulus pulses delivered to NAc during R- periods. Right: EMG potentials that triggered stimulation over each 20-hour session (gray), and their averages (black). Accepted biphasic patterns are followed by artifacts from triggered stimuli. **B.** Averages of biceps muscle activity surrounding NR-R schedule transitions (left) and R-NR transitions (right), shown separately for first third (blue) middle third (black) and last third (red) of the session. **C.** Peri-transition biceps activity during the 20-hour conditioning session, showing NR-R (left) and R-NR (right) transitions. Ordinates count the transitions over the course of the 20-hour session. The color indicates the rate of biceps EMG activity (see scale).

Figure 5. Response rates for cortical neurons during in-booth conditioning (A-C) and in-cage conditioning (D). A and B. initial and later sessions with monkey D. C. in-booth session with monkey J. D. in-cage session during free behavior with monkey J. Left: Points mark average rates during reinforced periods (red) when RCST stimulation was available and during non-reinforced periods (black). Bars denote 95% confidence intervals. Right. Statistics of cortical cell spike rate during R and NR periods shown as box plots. In each box, the central red line marks the distribution median, and blue box extremities depict upper (75%) and lower (25%) quartiles. Red crosses plot rates with values outside of whisker boundaries. Notch height shows approximate limits of confidence intervals about their median at the 5% significance level.

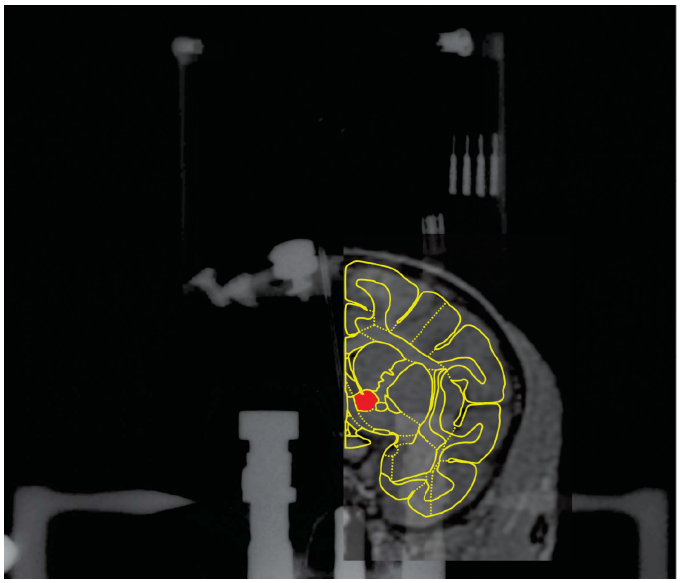
Figure 6. Peri-transition histograms of neuron spike activity during in-booth (A,B) and in-cage (C) conditioning. For each session, spikes occurring during two-minute intervals straddling NR-R transitions (left) and R-NR transitions (right) are pooled and binned into histograms above. Light-blue dashed vertical lines at t = 0 mark onset and offset of activity-dependent BSR. The point-density estimate of spike rate (thick red line) and its 95% confidence band limits (dashed red lines) overlay corresponding histograms. Horizontal gray lines show averages

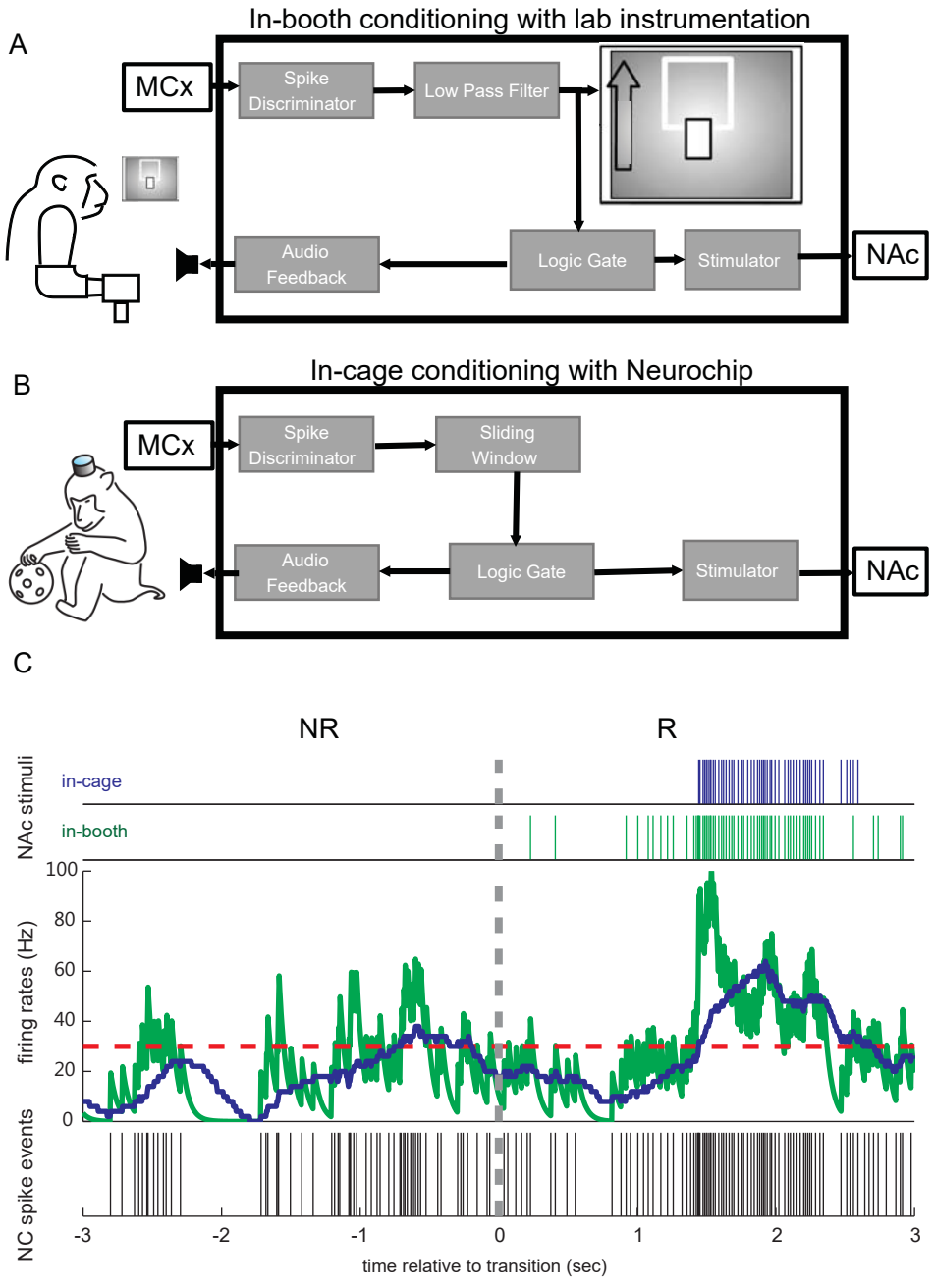
(solid) of sweeps, and surrounding 95% confidence intervals (dashed) after spike shuffling. Red line shows smoothed firing rates where the smoothing kernel width was determined by overall spike rate (Davison and Hinkley, 1997). The slow early rise in A is a result of low NR baseline rates followed by an abrupt increase. Bottom: averages of flexion-extension torques recorded concurrently. Depicted here are peri-transition activities from sessions: A. D2-2 and B. J:1-1. C. Peri-transition histograms of spike activity conditioned in-cage during session J:2-1. The NC2 delivered rate-contingent, spike-triggered BSR in the alternating R/NR schedule.

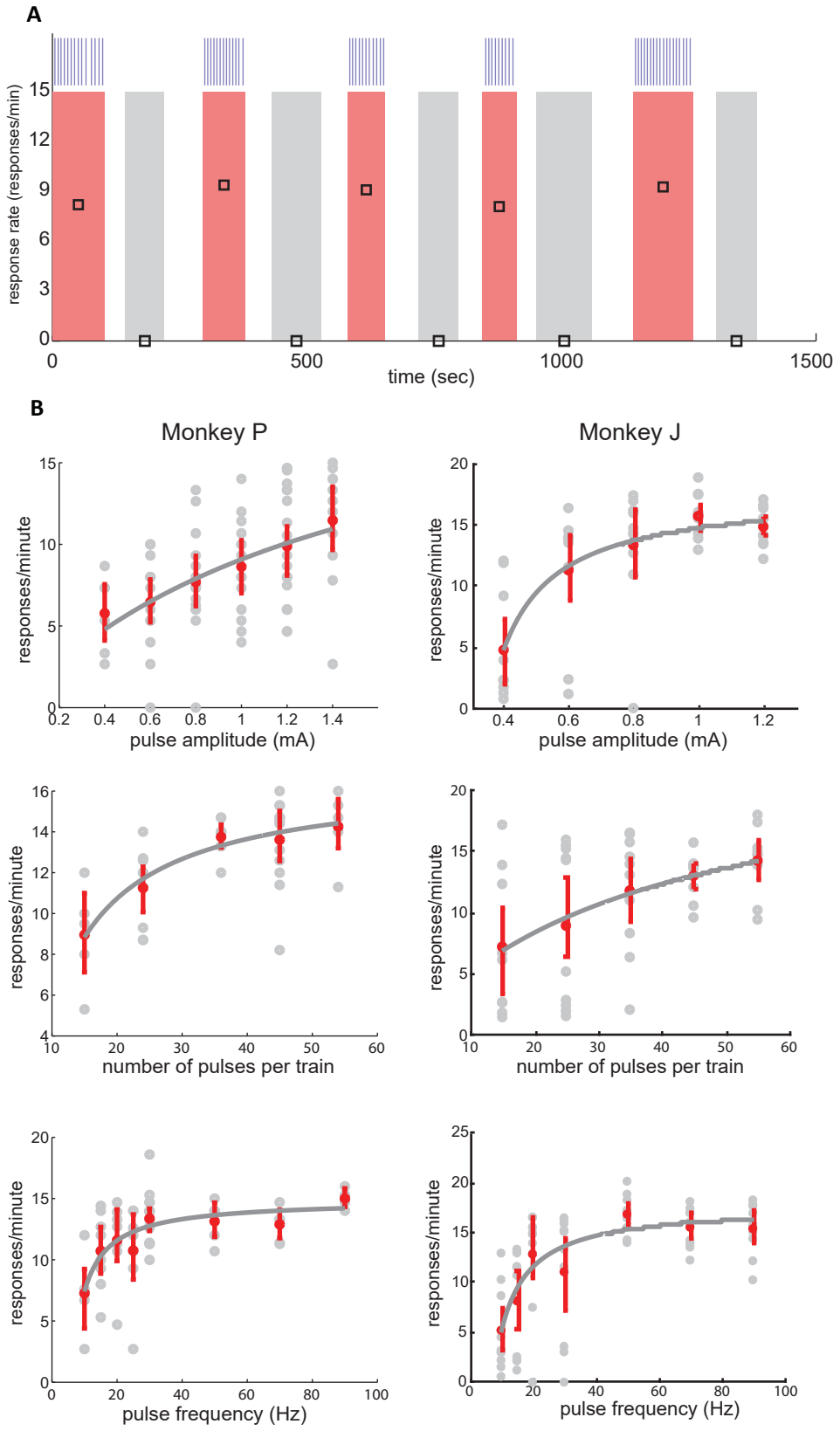
Figure 7. In-cage-conditioned spike activity grouped by first third (A), second third (B) and final third (C) of session. Spike activity from the example shown in Figure 6C.

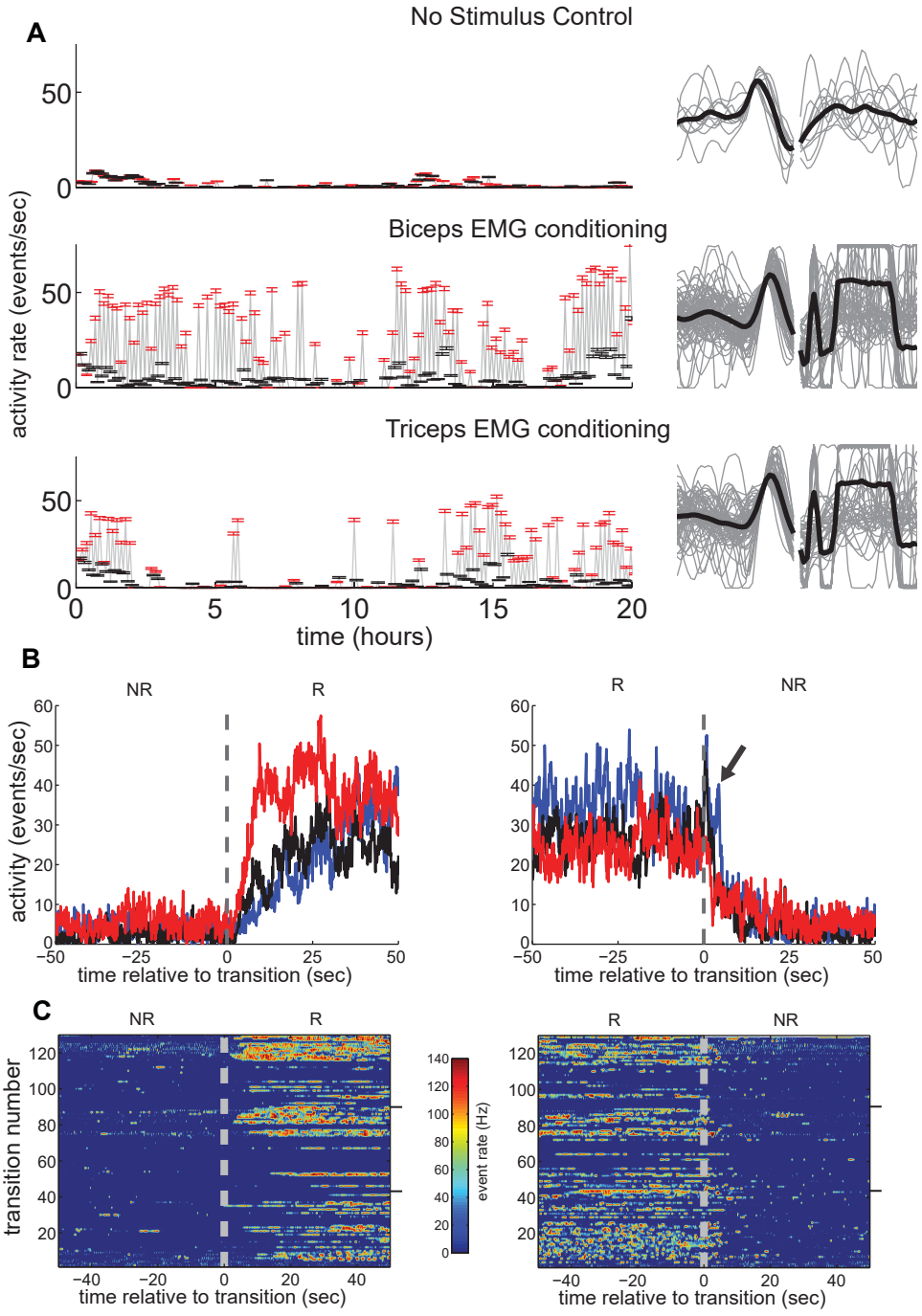
Figure 8. Rates of motor cortex neuron conditioned in-cage and in-booth with identical parameters. A. Average firing rates during R (red) and NR (black) and surrounding 95% confidence intervals. The monkey was in his home cage (light blue background) during the first hour and then quickly transferred to the training booth during second hour (tan). B. Peri-transition spike averages compiled during conditioning periods in-cage (left) and in-booth (right). Horizontal lines show overall mean rate for each environment. The Neurochip controlled conditioning in both environments; it ran continuously during the 6 minute transfer interval, and continued uninterrupted through the entire session. C. Statistics of cortical cell spike rate during R and NR periods shown as box plots for in-cage (left) and in-booth (right). The p values were obtained from the Kruskal-Wallis test.

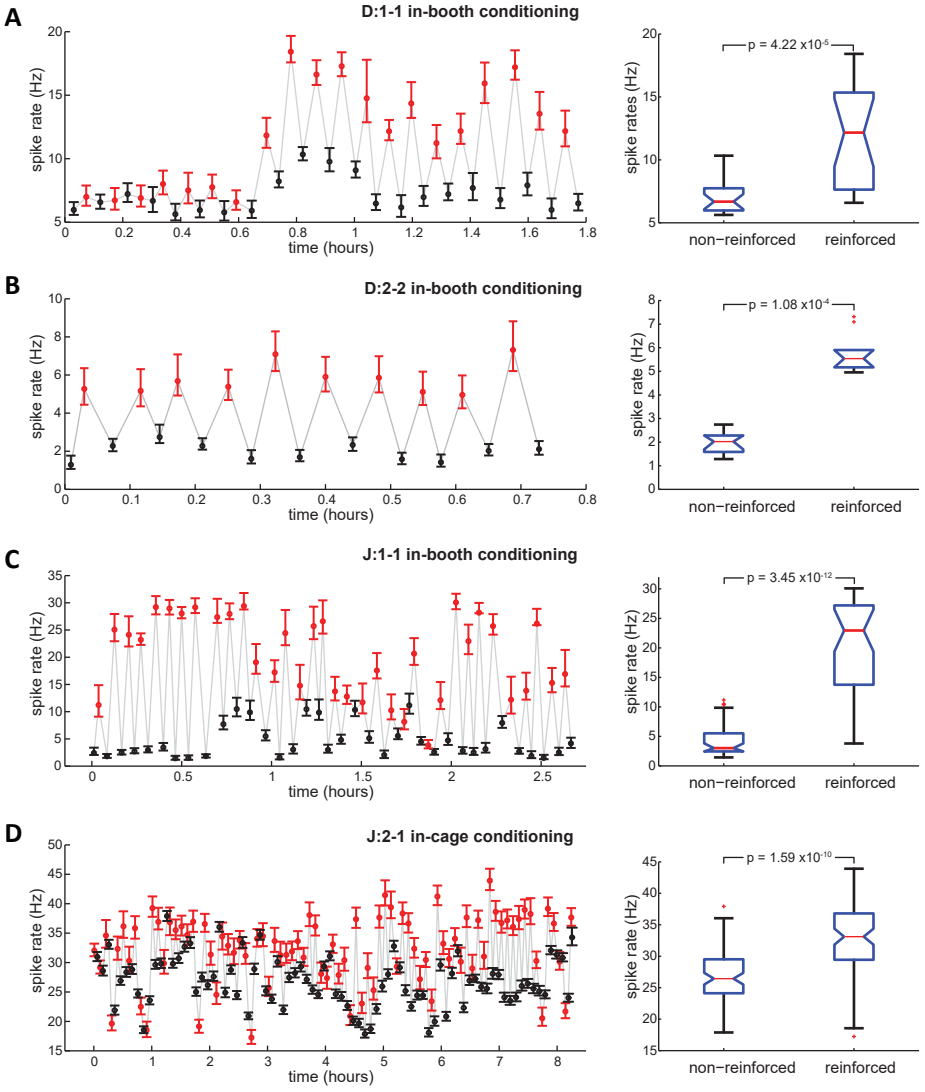
Figure 9. Peri-stimulus spike histograms compiled from spike trains surrounding single-pulse stimulation of NAc reinforcement sites. Histograms were compiled just prior to example experiments depicted in Figure 5A and Figure 5B using the same stimulation intensity. Vertical dashed lines depict stimulus delivery. Histogram bin width: 2ms. Point density average (thick red line) and limits of the 95% confidence band (red dashed lines) depict kernel-smoothed approximates of peri-stimulus mean spike rates.

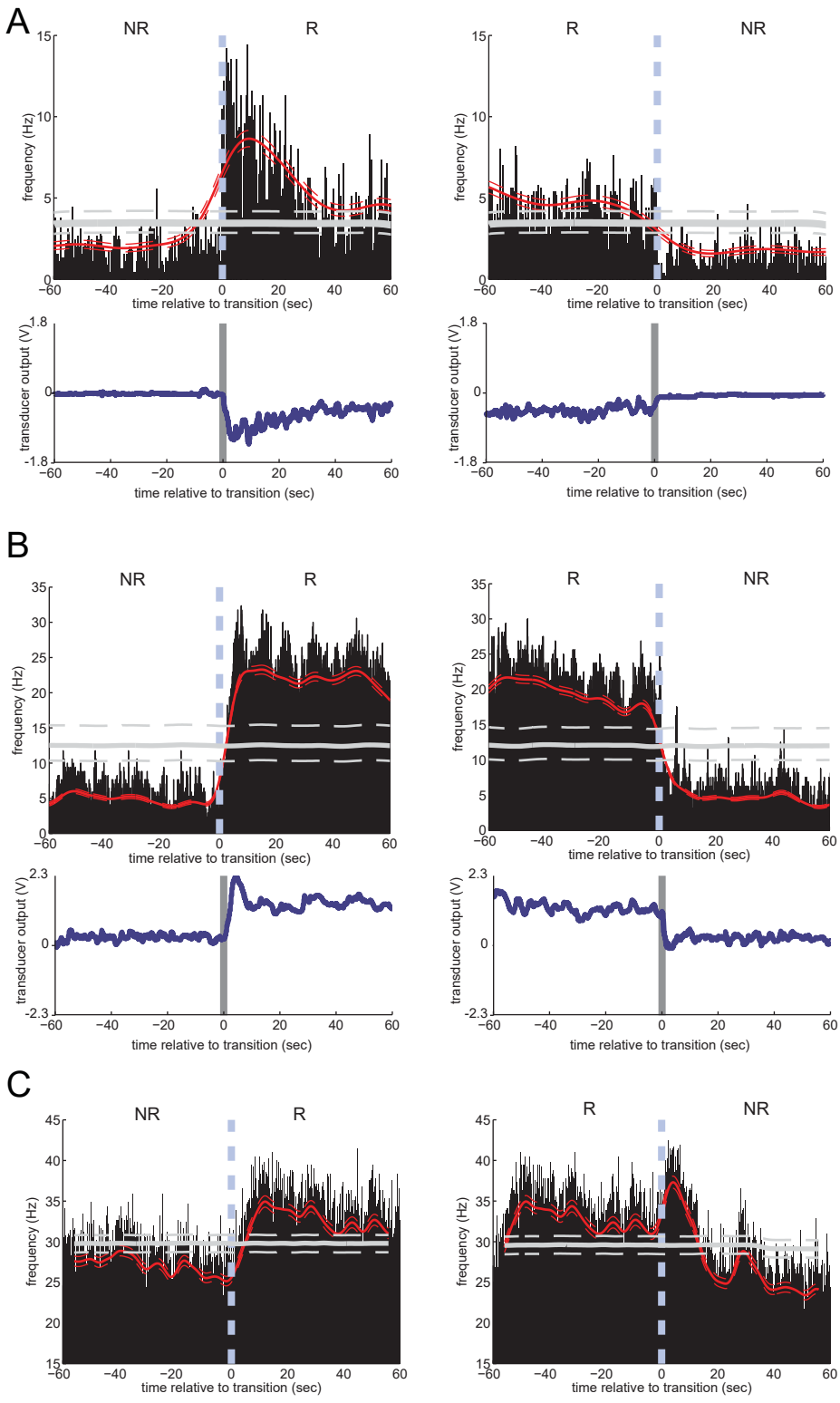


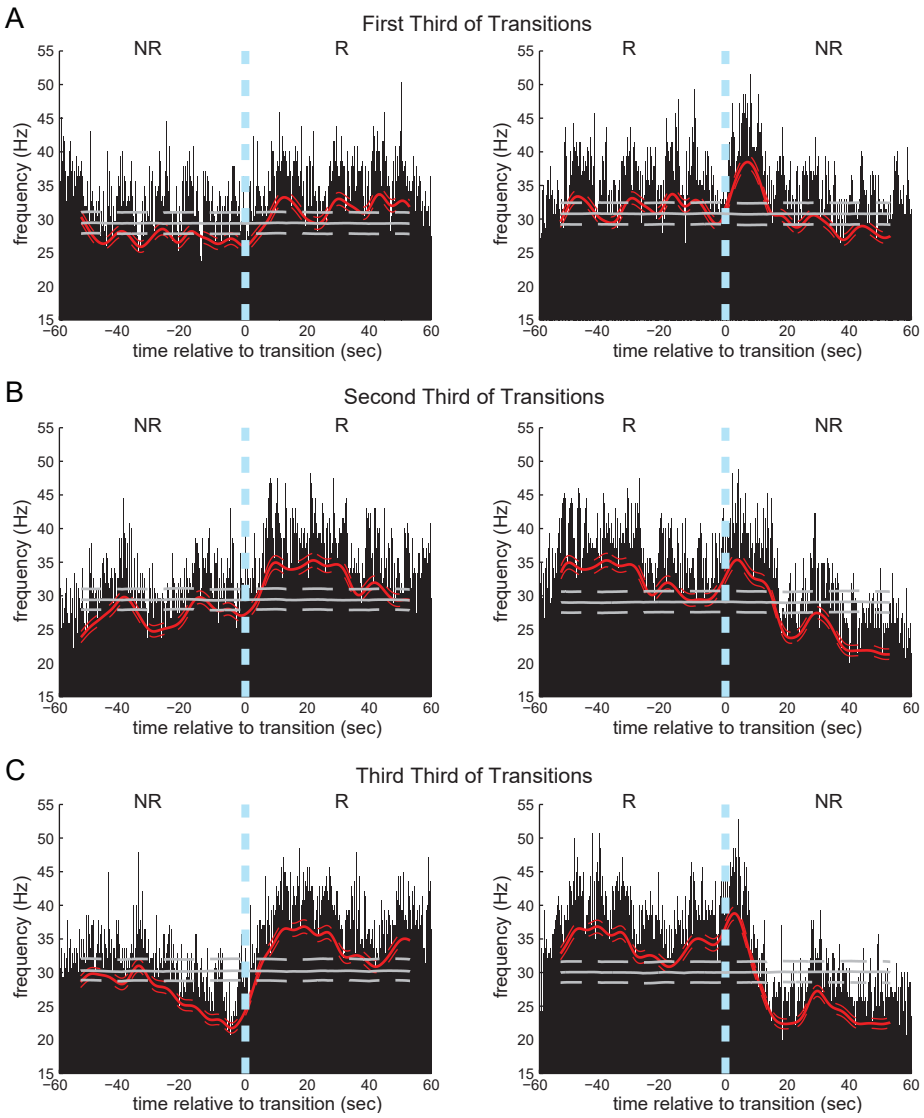

A
In-booth conditioning with lab instrumentation
MCx
Spike Discriminator
Low Pass Filter
Audio Feedback
Logic Gate
Stimulator
NAc
B
In-cage conditioning with Neurochip
MCx
Spike Discriminator
Sliding Window
Audio Feedback
Logic Gate
Stimulator
NAc
C
NR
R
NAc stimuli
in-cage
in-booth
firing rates (Hz)
100
80
60
40
20
0
NC spike events
-3
-2
-1
0
1
2
3
time relative to transition (sec)

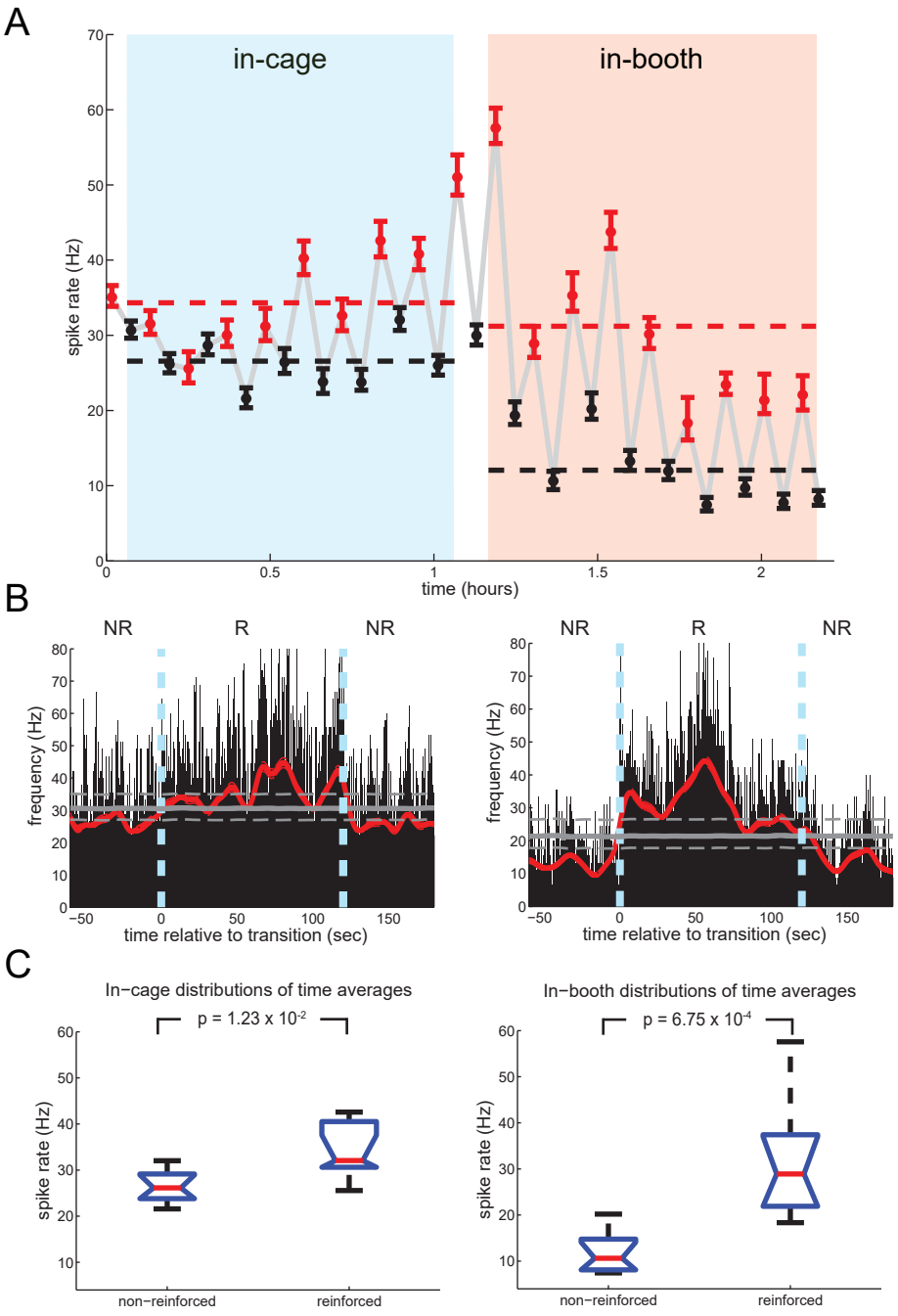

A
response rate (responses/min)
time (sec)
B
Monkey P
Monkey J
responses/minute
pulse amplitude (mA)
number of pulses per train
pulse frequency (Hz)

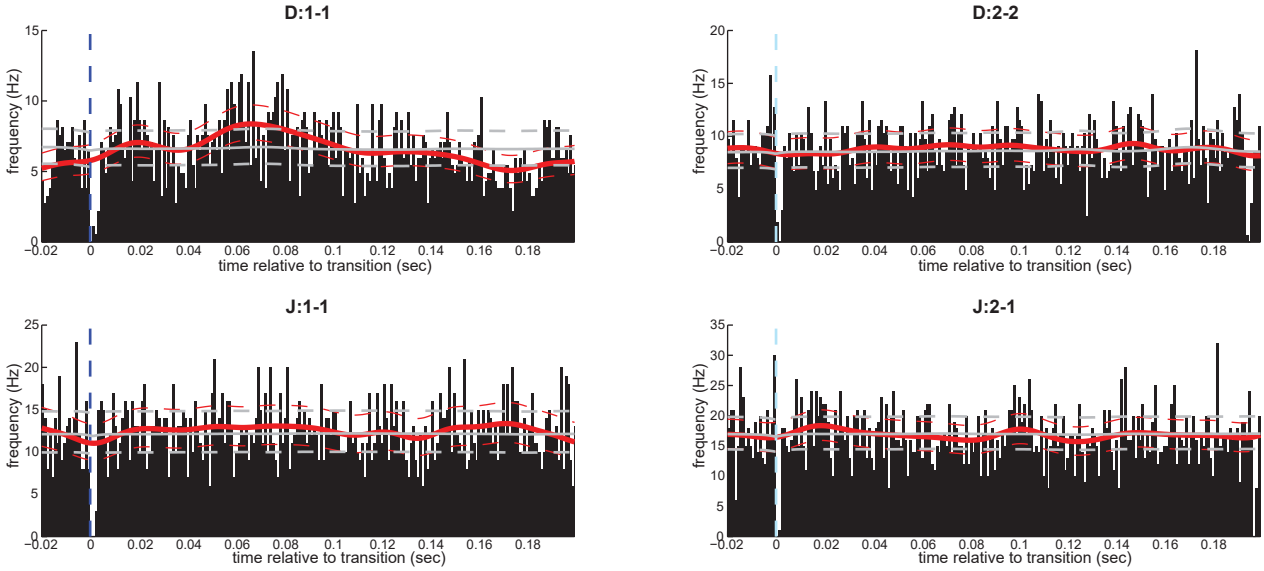

A
No Stimulus Control
Biceps EMG conditioning
Triceps EMG conditioning
activity rate (events/sec)
50
0
0
5
10
15
20
time (hours)
B
NR
R
activity (events/sec)
time relative to transition (sec)
C
transition number
event rate (Hz)
time relative to transition (sec)


A
D:1-1 in-booth conditioning
spike rate (Hz)
time (hours)
p = 4.22 x10-5
spike rates (Hz)
non−reinforced
reinforced
B
D:2-2 in-booth conditioning
spike rate (Hz)
time (hours)
p = 1.08 x10-4
spike rate (Hz)
non−reinforced
reinforced
C
J:1-1 in-booth conditioning
spike rate (Hz)
time (hours)
p = 3.45 x10-12
spike rate (Hz)
non−reinforced
reinforced
D
J:2-1 in-cage conditioning
spike rate (Hz)
time (hours)
p = 1.59 x10-10
spike rate (Hz)
non−reinforced
reinforced


A
NR
R
frequency (Hz)
time relative to transition (sec)
transducer output (V)
B
C


A
First Third of Transitions
NR
R
R
NR
frequency (Hz)
time relative to transition (sec)
B
Second Third of Transitions
C
Third Third of Transitions


A
in-cage
in-booth
spike rate (Hz)
time (hours)
B
NR
R
NR
frequency (Hz)
time relative to transition (sec)
C
In−cage distributions of time averages
p = 1.23 x 10-2
In−booth distributions of time averages
p = 6.75 x 10-4
spike rate (Hz)
non−reinforced
reinforced


D:1-1
D:2-2
J:1-1
J:2-1
frequency (Hz)
time relative to transition (sec)
−0.02
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18