Differential impact of reward and punishment on functional connectivity after skill learning

Adam Steel[1,2], Edward H. Silson[2], Charlotte J. Stagg[1,3,$], Chris I. Baker[2,$]

[1]Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, OX3 9DU, UK

[2]Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, 20814

[3]Oxford Centre for Human Brain Activity (OHBA), Wellcome Centre for Integrative Neuroimaging, University Department of Psychiatry, University of Oxford, Oxford, UK, OX3 9DU, UK

[$]Joint Senior Authors

Corresponding Author: Adam Steel, FMRIB Centre, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, United Kingdom; Tel. +44 (0) 1865 222493; email: adam.steel@nih.gov

Running title: Differential impact of reward and punishment on learning mechanisms

Number of figures: 6

Number of tables: 1

Abstract word count: 175

Introduction word count: 521

Discussion word count: 1585

Number of pages: 39

Competing financial interests: The authors declare no competing financial interests.

Acknowledgements: AS and CIB are funded by the NIMH internal research program (ZIA-MH002893). CJS holds a Sir Henry Dale Fellowship, funded by the Wellcome Trust and the Royal Society (102584/Z/13/Z). The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z). The authors would like to thank Matthew Rushworth for his helpful comments on the manuscript.

Abstract

Reward and punishment shape behavior, but the mechanisms underlying their effect on skill learning are not well understood. Here, we tested whether the functional connectivity of premotor cortex (PMC), a region known to be critical for learning of sequencing skills, is altered after training by reward or punishment given during training. Resting-state fMRI was collected in two experiments before and after participants trained on either a serial reaction time (n = 36) or force-tracking (n = 36) task with reward, punishment, or control feedback. In each experiment, training-related change in PMC functional connectivity was compared across feedback groups. In both tasks, reward and punishment differentially affected PMC functional connectivity. On the SRTT, participants trained with reward showed an increase in functional connectivity between PMC and cerebellum as well as PMC and striatum, while participants trained with punishment showed an increase in functional connectivity between PMC and medial temporal lobe connectivity. After training on the FTT, subjects trained with control and reward showed increases in PMC connectivity with parietal and temporal cortices after training, while subjects trained with punishment showed increased PMC connectivity with ventral striatum. While the results from the two experiments overlapped in some areas, including ventral pallidum, temporal lobe, and cerebellum, these regions showed diverging patterns of results across the two tasks for the different feedback conditions. These findings suggest that reward and punishment strongly influence spontaneous brain activity after training, and that the regions implicated depend on the task learned.

Keywords: Sequence learning, Motor learning, Consolidation, Motivation, Premotor cortex

The potential to use reward and punishment, collectively referred to as valenced feedback, during training has been pursued in recent years as a potential method to increase skill learning and retention (Wachter T et al. 2009; Abe M et al. 2011; Galea JM et al. 2015; Steel A et al. 2016). Prior behavioral studies of motor adaptation suggest that reward and punishment have differing effects on motor learning. For example, punishment increased learning rate in a cerebellar-dependent motor adaptation task (Galea JM *et al.* 2015), while reward prevented forgetting after adaptation (Shmuelof L et al. 2012; Galea JM *et al.* 2015). Reward may also restore adaptation learning in patients with cerebellar degeneration (Therrien AS et al. 2016) and stroke (Quattrocchi G et al. 2017). Beyond adaptation tasks, in other skill-learning contexts it has been reported that reward improves memory retention compared to punishment (Abe M *et al.* 2011), though these results are somewhat inconsistent across the literature (Steel A *et al.* 2016).

One potential explanation for the differential effects of reward and punishment on behavior is the recruitment of core set of brain regions involved in feedback processing. It has been suggested that punishment leads to the recruitment of fast learning systems [e.g. medial temporal lobe (MTL)], while reward recruits slow learning systems [e.g. caudate via dopaminergic signaling (Wachter T *et al.* 2009; Peterson EJ and CA Seger 2013)]. In support of this hypothesis, functional imaging studies have reported that reward increases caudate activity in a behaviorally-relevant manner (Wachter T *et al.* 2009; Peterson EJ and CA Seger 2013). In contrast, punishment increases activity in the anterior insula (Wachter T *et al.* 2009; Shigemune Y et al. 2014) and MTL (Murty VP et al. 2012; Murty VP, KS LaBar, et al. 2016). In studies of episodic memory, MTL activity during the post-encoding period subserves the reward-related memory improvement (Gruber MJ et al. 2016; Murty VP, A Tompary, et al. 2016), but the relevance of the post-encoding period has not been studied in a skill learning context, nor have the differential effects of training with reward and punishment on post-encoding neural activity been studied.

In the present study, we tested whether reward and punishment affect neural activity of the slow and fast learning systems immediately after skill-learning, respectively (Doyon J et al. 2018). In the first experiment, based on the well described role of fast and slow learning systems in the context of the serial reaction time task [SRTT; (Doyon J *et al.* 2018)], we examined the impact of valenced feedback on neural activity induced by SRTT training. In the second experiment, we implemented the force tracking task (FTT) with reward and punishment to test whether any regions impacted by training with feedback generalize to a different motor sequencing task with distinct task demands. In both experiments, before and after training we collected 20-minutes of resting-state fMRI data (Figure 1a-d). We have previously presented our behavioral results, which suggested that feedback differentially impacted on performance during learning in the two tasks and that all groups showed equal retention (Steel A *et al.* 2016). We wished to extend these findings by analyzing brain activity change due to training. To that end, we examined the change in functional connectivity after training on each task.

We focused on premotor cortex (PMC) as the key region for evaluating post-encoding connectivity in both tasks, based on its well-documented critical role as a memory-encoding region for sequence and sensorimotor learning (Floyer-Lea A and PM Matthews 2004, 2005; Hardwick RM et al. 2013; Wiestler T and J Diedrichsen 2013; Kornysheva K and J Diedrichsen 2014). In addition, PMC shows reward-related activity after movement (Ramkumar P et al. 2016). We predicted that PMC functional connectivity would be differentially modulated by training with reward and punishment, with connectivity between the PMC and the anterior insula, MTL, cerebellum, and caudate providing distinct contributions to skill retention. We found that reward and punishment did differentially impact PMC functional connectivity after training, but that task demands modulated the impact of feedback; after training on the SRTT, participants trained with reward and punishment showed differentially functional connectivity between PMC with the striatum, cerebellum, and MTL. However, after training on the FTT, participants trained with reward and punishment showed differential functional connectivity between PMC with frontal and parietal areas. These results suggest that feedback valence strongly affects neural processing after learning, and that the influence of reward and

punishment is limited to regions involved in task performance rather than a core set of feedback processing hubs.

Materials and methods

Overview

In two experiments, participants were trained on either the serial reaction time task (SRTT) or the force-tracking task (FTT) with reward, punishment, or uninformative feedback (Figure 1A). No participant was trained in both tasks. A detailed description of the tasks and training procedure can be found in (Steel A *et al.* 2016) and is summarized below. Before and after the training session, 20-minutes of resting-state fMRI was collected.

Participants

78 participants (47 female, mean age = 25 years ± std. 4.25) were recruited and participated in the study. All participants were right-handed, free from neurological disorders, and had normal or corrected-to-normal vision. All participants gave informed consent and the study was performed with National Institutes of Health Institutional Review Board approval in accordance with the Declaration of Helsinki (93-M-0170, NCT00001360). Data from six individuals were removed from the study due to inattention during training (defined as non-responsive or inaccurate performance on greater than 50% of trials; n=3) or inability to complete the imaging session due to discomfort or fatigue (n=3). This left 72 participants with complete data sets included in the analyses presented here.

Training procedure

Both tasks followed the same behavioral training procedure. Trials were presented over 15 blocks with a 30-second break separating each. Unbeknownst to the participants, during some

blocks ("fixed-sequence blocks") the stimulus would appear according to a repeating pattern (described below for each task). During other blocks the appearance of the stimulus was randomly determined ("random-sequence blocks").

To familiarize participants to the task, and establish their baseline level of performance, the task began with three random-sequence blocks without feedback ("familiarization blocks"). Participants were unaware of the forthcoming feedback manipulation during the familiarization blocks. Then the feedback period began, starting with a pre-training probe (three blocks, random – fixed – random), then the training blocks (six consecutive fixed-sequence blocks), and, finally, a post-training probe (three blocks, random – fixed – random). The difference in performance between the mean of the two random blocks compared to the fixed sequence block, during these probes was used to index sequence knowledge (Robertson EM 2007). Participants were presented with only one sequence during the fixed-sequence blocks.

To test the impact of reward and punishment on skill learning, participants were randomised into one of 3 feedback groups: reward, punishment, or uninformative (control). During the feedback period, reward, punishment, or control feedback was provided based on the participant's ongoing performance. The feedback paradigm for each task is outlined separately below.

Training was conducted inside the MRI scanner, and functional MR images were collected during the training period. These training period data are outside the scope of the present manuscript and are not presented here.

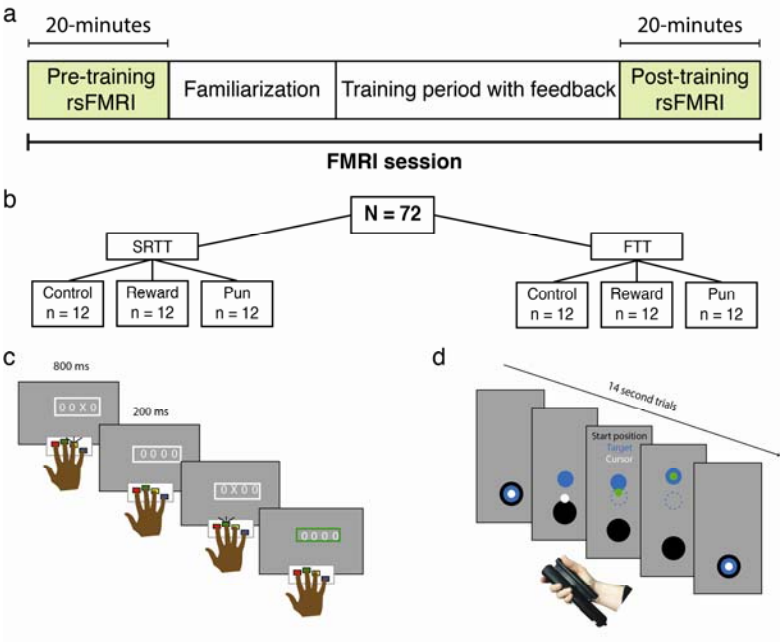

Figure 1. Experimental design and skill memory retention. (a,b) Participants underwent 20 minutes of resting state fMRI before and after training on either the serial reaction time task (SRTT) or the force tracking task (FTT) while receiving REW, PUN, or CONT feedback. In the SRTT (c), participants responded to a cue appearing in one of four locations on a screen. In the FTT (d), participants modulated their grip force to track a moving target. In both tasks, the stimulus could follow either a random or fixed sequence, and skill memory was assessed by comparing performance during random and fixed trials.

Experiment 1: Serial reaction time task (SRTT)

The version of the SRTT used here adds feedback to the traditional implementation. At the beginning of each block participants were presented with four "O"s, arranged in a line, at the center of the screen. These stimuli were presented in white on a grey background (Figure 1B). A trial began when one of the "O"s changed to an "X". Participants were instructed to respond as quickly and accurately as possible, using the corresponding button, on a four-button response device held in their right hand. The "X" remained on screen for 800 ms regardless of whether the subject made a response, followed by a 200 ms fixed inter-trial interval, during

which time the four "O"s were displayed. While this trial timing may foster some degree of explicit awareness in some subjects, making this variant of the SRTT not a purely motor learning task, this timing was necessary to accommodate the constraints of collecting fMRI data during training.

A block consisted of 96 trials. During fixed-sequence blocks, the stimuli appeared according to one-of-four fixed 12-item sequences, which repeated 8 times (e.g. 3-4-1-2-3-1-4-3-2-4-2-1). For each participant, the same 12-item sequence was used for the duration of the experiment. Each fixed block began at a unique position within the sequence, to help prevent explicit knowledge of the sequence from developing (Schendan HE et al. 2003). In the random blocks, the stimuli appeared according to a randomly generated sequence, without repeats on back-to-back trials, so, for example, subjects would never see the triplet 1-1-2.

Breaks between blocks lasted 30-seconds. Initially, participants saw the phrase "Nice job, take a breather". After five seconds, a black fixation-cross appeared on the screen. Five seconds before the next block began, the cross turned blue to cue the subjects that the next block was about to start.

During the post-training retention probes, participants performed three blocks (random – fixed – random), outside the scanner on a 15-inch Macbook Pro using a button box identical to the one used during training. During these retention probes, the next trial began 200 ms after the participant initiated their response rather than after a fixed 800 ms as during training. No feedback was given during the retention blocks.

Experiment 2: Force-tracking task

In the force-tracking task (FTT), participants continuously modulated their grip force to match a target force output (Floyer-Lea A and PM Matthews 2005; Floyer-Lea A et al. 2006). In the traditional implementation, participants are exposed to a single pattern of force modulation

repeated each trial. This design does not allow discrimination between general improvement (i.e. familiarization with the task and/or the force transducer) and improvement specific to the trained sequence of force modulation. Therefore, we adapted the traditional FTT method to align it with the experimental design that is traditional for the SRTT, i.e. by including random sequence blocks.

A given trial consisted of a 14 second continuous pattern of grip modulation. At the beginning of a trial, participants were presented with three circles on a grey background projected onto a screen: a white circle (Cursor, 0.5 cm diameter), a blue circle (Target, 1.5 cm diameter), and a black circle (Bottom of the screen, 2 cm diameter, indicating the position corresponding to minimum pressure; Figure 1C). Participants held the force transducer (Current Designs, Inc., Philadelphia, PA) in the right hand between the four fingers and palm (Figure 1D, bottom). Participants were instructed to squeeze the force transducer (increasing force moving the cursor upwards) to keep the cursor as close to the center of the target as possible as the target moved vertically on the screen. During fixed blocks, participants were randomly assigned to one of six sequences (Figure 1D, left). During random blocks, the target followed a trajectory generated by the linear combination of four waveforms, with periods between 0.01 and 3 Hz. The combinations of waveforms were constrained to have identical average amplitude (target height), and the number and value of local maxima and minima were constant across the random blocks.

For data analysis, the squared distance from the cursor to the target was calculated at each frame refresh (60 Hz). The first 10 frames were removed from each trial. The mean of the remaining time points was calculated to determine performance, and trials were averaged across blocks.

Feedback

All participants were paid a base remuneration of $80 for participating in the study. At the start of the feedback period, participants were informed they could earn additional money based on their performance.

For full details of our tasks please see (Steel et al, 2016a). In the SRTT, performance was defined as the accuracy (correct or incorrect) and reaction time (RT) of a given trial. Feedback was given on a trial-by-trial basis (Figure 1C,D). This was indicated to the participant when the white frame around the stimulus changed to green (reward) or red (punishment). In the reward group, the participants were given feedback if their response was accurate and their RT was faster than their criterion RT, which indicated that they earned money ($0.05 from a starting point of $0) on that trial. In the punishment group, participants were given feedback if they were incorrect, or their RT was slower than their criterion, which indicated that they lost money ($0.05 deducted from a starting point of $55) on that trial. Participants in the control-reward and control-punishment groups saw red or green color changes, respectively, at a frequency matched to punishment and reward, respectively. Control participants were told that they would be paid based on their speed and accuracy. Importantly, to control for the motivational differences between gain and loss, participants were not told the precise value of a given trial. This allowed us to assess the hedonic value of the feedback, rather than the level on a perceived-value function. Between blocks, for the reward and punishment groups, the current earning total was displayed (e.g. "You have earned $5.00"). Control participants saw the phrase, "You have earned money." The criterion RT was calculated as median performance in the first familiarization block. After each block, the median + standard deviation of performance was calculated, and compared with the criterion. If this test criterion was faster (SRTT) or more accurate (FTT) than the previous criterion, the criterion was updated. During the SRTT, only the correct responses were considered when establishing the criterion reaction time.

Feedback in the FTT was based on the distance of the cursor from the target (Figure 1C). For the reward group, participants began with $0. As participants performed the task, their cursor turned from white to green when the distance from the target was less than their criterion.

This indicated that they were gaining money at that time. In the punishment group, participants began with $45, and the cursor turned red if it was outside their criterion distance. This indicated that they were losing money. For reward-control and punishment control, the cursor changed to green or red, respectively, but was unrelated to their performance. For control, the duration of each feedback instance, as well as cumulative feedback given on each trial, was matched to the appropriate group. Consistent with the SRTT experiment, between each block, participants were shown their cumulative earnings. Control participants saw the phrase "You have money."

MRI acquisition

This experiment was performed on a 3.0T GE 750 MRI scanner using a 32-channel head coil (GE Medical Systems, Milwaukee, WI).

Structural scan

For registration purposes, a T1-weighted anatomical image was acquired (magnetization-prepared rapid gradient echo (MPRAGE), TR = 7 ms, TE = 3.4 ms, flip-angle = 7 degrees, bandwidth = 25.000 kHz, FOV = 24x24 cm^2, acquisition matrix = 256 x 256, resolution = 1 x 1 x 1 mm, 198 slices per volume). Grey matter, white matter, and CSF maps for each participant were generated using Freesurfer (Fischl B et al. 2002).

EPI scans

Both task and resting state fMRI scans were collected with identical parameters and slice prescriptions. Multi-echo EPI scans were collected with the following parameters: TE = 14.9, 28.4, 41.9 ms, TR = 2, ASSET acceleration factor = 2, flip-angle = 65 degrees, bandwidth = 250.000 kHz, FOV = 24 x 24 cm, acquisition matrix = 64 x 64, resolution = 3.4 x 3.4 x 3.4 mm, slice gap = 0.3 mm, 34 slices per volume covering the whole brain. Respiratory and cardiac

traces were recorded. Each resting state scan lasted 21-minutes. The first 30 volumes of each resting-state scan were discarded to control for the difference in arousal that occurs at the beginning of resting state scans. This left the final 20-minutes of rest in each scan for our analysis. This procedure has been used in other studies where long-duration resting state runs were collected (Gonzalez-Castillo J et al. 2014).

Resting state fMRI preprocessing

Data were preprocessed using AFNI (Cox RW 1996). The time series for each TE was processed independently prior to optimal combination (see below). Slice-time correction was applied (3dTShift) and signal outliers were attenuated [3dDespike (Jo HJ et al. 2013)]. Motion correction parameters were estimated relative to the first volume of the middle TE (28.4 msec), and registered to the structural scan (3dSkullStrip, 3dAllineate). These registration parameters were then applied in one step (3dAllineate) and the data were resampled to 3 mm isotropic resolution.

The optimal echo time for imaging the BOLD effect is where the TE is equal to T2*. Because T2* varies across the brain, single echo images are not optimal to see this variation. By acquiring multiple echoes, this enables the calculation of the "optimal" T2* weighted average of the echoes, which allows one to recover signals in dropout areas and improves contrast-to-noise ratio (Posse S et al. 1999; Poser BA et al. 2006; Kundu P et al. 2014; Evans JW et al. 2015). The following is a summary of methods implemented in the ME-ICA procedure.

The signal at an echo, *n* varies as a function of the initial signal intensity S_0 and the transverse susceptibility T2* = 1/R2* and is given by the mono-tonic exponential decay:

$$S(TE_n) = S_0 exp(-R_2 * TE_n),$$

where R2* is the inverse of relaxation time or 1/T2*. This equation can be linearized to simplify estimation of T2* and S_0 as the slope using log-linear transformation. The time courses can be optimally combined by weighted summation by a factor, w, described by the following equation:

$$w(T^*_{2(fit)})_n = \frac{TE_n \cdot exp(-TE_n/T^*_{2(fit)})}{\sum_n \; TE_n \cdot exp(-TE_n/T^*_{2(fit)})}.$$

Where T_2(fit) is the transverse relaxation time estimated for each voxel using the equation above. The OC time series can then be treated as a single echo, as it is here for the resting state data.

After optimal combination, we applied the basic ANATICOR (Jo HJ et al. 2010) procedure to yield nuisance time series for the ventricles and local estimates of the BOLD signal in white matter. All nuisance time-series (six parameters of motion, local white matter, ventricle signal, and 6 physiological noise regressors (AFNI: RetroTS)) were detrended with fourth order polynomials. These time series, along with a series of sine and cosine functions to remove all frequencies outside the range (0.01-0.25 Hz) were regressed out in a single regression step (AFNI program 3dTproject). Time points with motion greater than 0.3 mm were removed from the data [scrubbing, see Power JD et al. (2012)] and replaced with values obtained via linear interpolation in time. Data were transformed into standard space (@auto_tlrc) and smoothed with a 6mm FWHM Gaussian kernel. For group data analysis and calculation of global connectedness, a group-level grey matter by mask was created by calculating voxels in standard space determined to be grey matter in 80% of participants (Gotts SJ et al. 2012) fMRI data analysis

For group tests, the average smoothness of the data was estimated (3dFWHMx). Data were corrected for multiple comparisons using Monte-Carlo simulations to (3dClustSim, AFNI compile date July 9, 2016). Cluster size correction was applied to achieve an $\alpha = 0.05$ ($p < 0.005$, $k = 55$) unless otherwise indicated.

Left premotor cortex functional connectivity

We focused our analysis on the left premotor cortex based on this regions well-described central role in sequence production in response to visual cues (Mushiake H et al. 1991), sensorimotor learning, and sequence learning (Hardwick RM *et al.* 2013; Hardwick RM et al. 2015). Given that the participants were performing the task with their right hand, we further focused on the left PMC. Left dorsal and ventral premotor cortex (PMd and PMv) were defined based on a publicly available diffusion-MRI based parcellation of premotor cortex (Tomassini V et al. 2007), which are matched for size. The mean time series from both dorsal and ventral premotor cortices were extracted separately from each participant and each rest period, and the whole brain correlation maps (Pearson's r) for both PMd and PMv during the pre- and post-training resting state MRI scans were then calculated based on these time series. For each resting state scan, global correlation [GCOR; @compute_gcor: Saad ZS et al. (2013)] and magnitude of motion across runs (@1dDiffMag) were calculated and included as nuisance covariates for group analysis.

The resulting maps were then submitted to a linear mixed effects model (3dLME) with ROI (PMd/PMv), Rest (pre-/post-), Group (Control/Reward/Punishment). The precise model fit at the group level in R-syntax (nlme) was 'Group*Rest*ROI+motion+gcor'. Group analysis maps were cluster-corrected for multiple comparisons to achieve a $\alpha = 0.05$ using the ACF model in 3dClustSim ($p < 0.005$, $k = 54$).

To determine the consistency of the impact of feedback across the two tasks, the overlap of the significant clusters from the Rest x Condition interaction for the SRTT and FTT was calculated at both a conservative ($p < 0.005$, $k = 54$) and liberal ($p < 0.01$, $k = 54$) threshold.

Results

The behavioral data have previously been reported [Steel A *et al.* (2016)]. Briefly, during performance of the SRTT, we found that, compared to participants training with reward, those training with punishment showed reduced reaction time without detriment to accuracy; in contrast, on the FTT, compared to participants training with reward, those training with punishment exhibited greater tracking error during training. During retention testing, in both tasks all feedback groups showed evidence of sequence knowledge at 1hr, 24-48hrs and 30-d after training. To identify brain regions impacted by feedback given during training, we used a seed-based analysis focused on the left PMC. For each experiment, we implemented a voxel-wise linear mixed effects (LME) model (Chen G et al. 2013) with Rest (Pre-/Post- training), Group (Reward /Punishment /Control), and ROI (ventral-/dorsal- premotor cortex) as factors and looked for regions showing a significant interaction between Rest x Condition. No regions showed a Rest x Condition x ROI interaction in either task. Full results from this model for both tasks can be found in Table 1.

Reward and punishment evoke dissociable changes in PMC connectivity patterns after SRTT training

In Experiment 1, following training on the SRTT task, several regions showed a modulation of PMC functional connectivity that differed by the feedback given during training (Rest x Group interaction; Figure 2). These included bilateral thalamus and striatum, right cerebellar vermis, supplementary motor area (SMA), bilateral medial temporal lobe, and left inferior frontal gyrus. In order to examine the nature of the interaction, we extracted the estimated mean PMC

functional connectivity change due to training from each cluster (Figure 2, inset). This revealed a pattern that was distinct across the feedback groups. After training with reward, functional connectivity increased between PMC and thalamus and striatum, cerebellar vermis, and SMA, but PMC functional connectivity decreased with medial temporal lobe and inferior frontal gyrus. The punishment group showed the opposite pattern; PMC connectivity with medial temporal lobe and inferior frontal gyrus increased after training with punishment. The control group showed in intermediate pattern between the reward and punishment groups. In the control group, PMC connectivity increased with the thalamus, striatum, and cerebellar vermis after training, but decreased with SMA, medial temporal lobe, and inferior frontal gyrus. Thus, in the context of the SRTT, there is evidence multiple streams can support memory formation, which can be dissociated by augmenting training with reward or punishment.

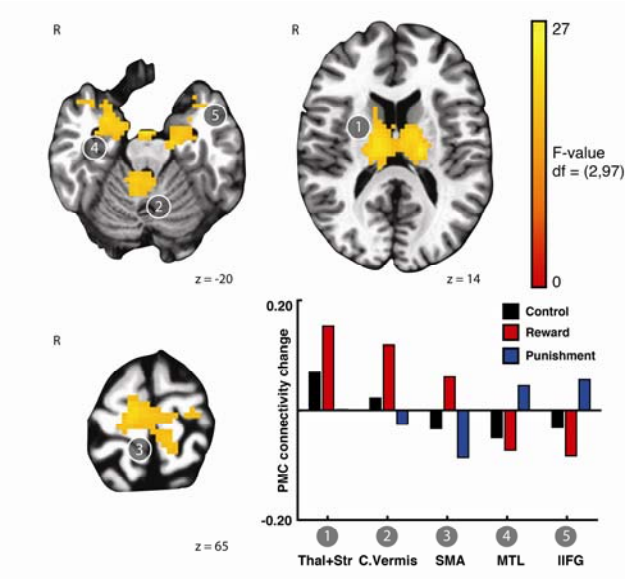

Figure 2. Reward and punishment differentially effect PMC functional connectivity change after training on the SRTT. Linear mixed effects model revealed brain regions exhibiting a Rest x Feedback group interaction in the functional connectivity of left PMC. Functional connectivity between PMC and the thalamus and striatum, cerebellum, and SMA increased after training with reward and control feedback but decreased after training with punishment. In contrast, functional connectivity between medial temporal lobe and left

inferior frontal gyrus increased after training with punishment but decreased after training with reward and control feedback. Bars show mean connectivity change after training estimated by the linear mixed effects model.

Punishment promotes PMC-striatal connectivity after FTT training

In Experiment 2, after training on the FTT, PMC connectivity change with bilateral lateral occipital temporal cortex, precuneus, left culmen, cerebellum, right striatum, and right medial frontal gyrus differed across the feedback groups (Figure 3). Extracting the estimated mean PMC functional connectivity change of these regions revealed that two primary modes of change in these regions (Figure 3, inset). PMC connectivity with the lateral occipital cortex, precuneus, culmen, and cerebellar lobule VI, decreased after training with punishment, but increased after training with control feedback. In contrast, connectivity of the PMC with right striatum and right medial frontal gyrus increased after training with punishment. PMC connectivity to right medial frontal gyrus decreased after training with control. Overall, in the regions showing the interaction between functional connectivity change after training and feedback group, reward had a smaller influence compared to control or punishment. This suggests that punishment induces marked changes in functional connectivity change after FTT training, which was consistent with the behavioral effects observed at the end of training (Steel A *et al.* 2016).

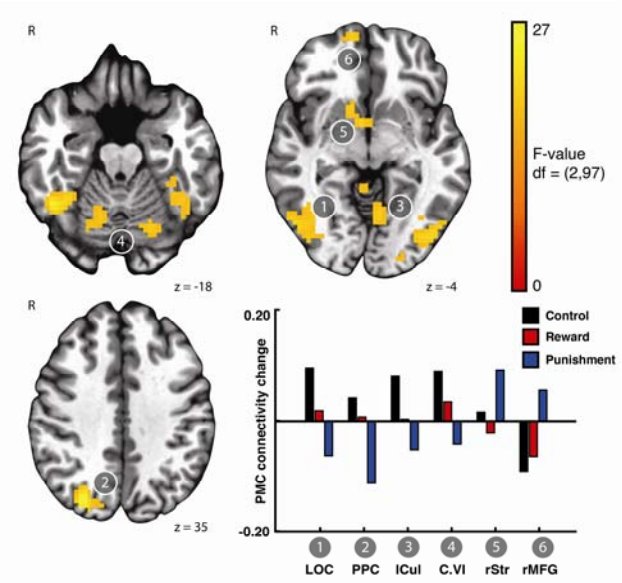

Figure 3. Punishment promotes PMC-striatal connectivity after FTT training. Linear mixed effects model revealed brain regions exhibiting a Rest x Feedback group interaction in the functional connectivity of left PMC. Functional connectivity between the PMC and lateral occipital cortex, posterior parietal cortex, culmen, and cerebellum increased after training with control feedback, but decreased after training with punishment. In contrast, PMC connectivity with Right ventral striatum and right medial frontal gyrus increased after training with punishment but decreased after training with reward and control feedback. Bars show mean connectivity change after training estimated by the linear mixed effects model.

Influence of reward and punishment on PMC functional connectivity depends on task

In order to better understand the correspondence between the feedback and the change in PMC functional connectivity after training between the two tasks, we calculated the overlap between the sets of regions that showed the Rest x Group interaction for each task separately. Given the reduced power to detect general effects that might be weak but consistent across tasks, for this qualitative assessment, we considered the conservative threshold reported above

($p < 0.005$, $k = 54$), as well as a liberal threshold ($p < 0.01$, $k = 54$; Figure 4a). The mean connectivity change with PMC for clusters greater than 20 voxels in extent is shown in Figure 4b. At the conservative threshold, one region in the striatum showed overlap (Cluster 1). When we extracted the estimated mean connectivity change after training for each task, however, we found that the pattern of connectivity change across the feedback groups was diametrically opposed. In the SRTT, connectivity between the PMC and the overlapping region increased after training with reward but decreased after training with punishment. In contrast, after training on the FTT, connectivity between PMC and the overlapping region increased after training with punishment but increased after training with reward. In both tasks, connectivity between PMC and the overlapping region increased after training with control feedback.

At the liberal threshold, we detected three other clusters in the right anterior temporal lobe, the right cerebellum, and right dorsal thalamus. In each of these clusters, the pattern of connectivity change with PMC exhibited by the Reward and Punishment groups was also not consistent across tasks.

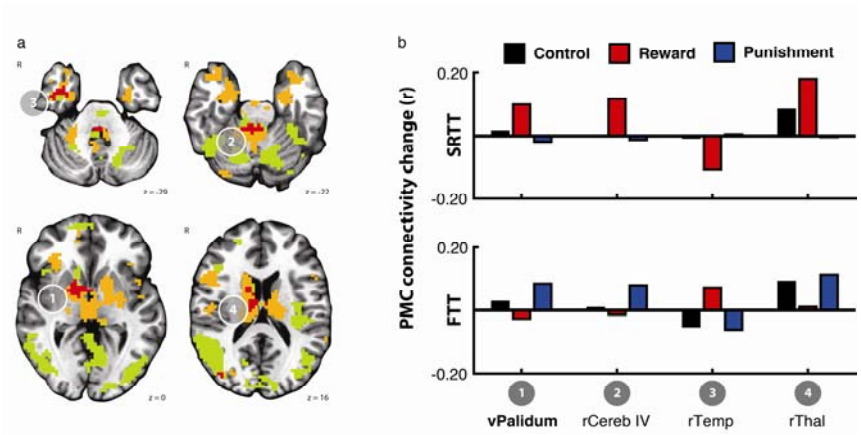

Figure 4. Regions showing PMC functional connectivity change after training were had limited overlap across tasks. Clusters that showed a PMC connectivity change after training that differed by the feedback group in the SRTT (orange) and the FTT (green) showed limited overlap, even at a liberal cluster threshold ($p < 0.01$, $k = 54$). Only the ventral pallidum

(Cluster 1, bold) showed an overlap at the threshold reported above ($p < 0.005$, $k = 54$). In all clusters, the Reward and Punishment groups exhibited opposite patterns of connectivity change with PMC due to training on the SRTT (upper) and FTT (lower). Bars show mean connectivity change after training estimated by the linear mixed effects model.

To directly compare the effect of task on the functional connectivity change after training with feedback, we performed a whole-brain analysis (cluster corrected; $p < 0.005$, $k = 54$) using a linear mixed effects model that included Task (SRTT/FTT), Rest (Pre-/Post-training), and Feedback group (Control/Reward/Punishment) as factors; for this analysis, we specifically focused on regions that showed a Task x Feedback x Rest interaction. Critically, all regions that appeared in the overlap analysis reported above evidenced a Task x Feedback x Rest interaction (bilateral basal ganglia including caudate, putamen, and pallidum [peak x=-10, y=16, z=17], right cerebellum [1.5, 24.5, -24.5], right medial temporal lobe [-40, 4.5, -33.5]). In addition the right lateral temporal lobe [-43, 49, -15], right parietal cortex [-43, 55, 17.5], right medial frontal gyrus [-13.5, 52.5, -6.5], and right lateral prefrontal cortex [-46.5 -16.5 17.5] also showed this interaction (Figure S1).

In summary, it is clear that feedback profoundly impacts resting state functional connectivity changes after training. However, similar to behavioral performance, the impact regions of the brain impacted by feedback depend critically on the task being performed.

Discussion

The present study sought to investigate the effect of valenced feedback on brain activity during the period immediately after training on the SRTT and FTT. We found that reward and punishment differentially impacted change in PMC functional connectivity, and this impact was influenced by task. After training on the SRTT, functional connectivity between PMC with striatum, cerebellum, and SMA increased after training with reward, while functional

connectivity between PMC connectivity with medial temporal lobe and inferior frontal gyrus increased after training with punishment. Training on the FTT induced a different pattern of connectivity change. After training on the FTT, functional connectivity between PMC and parietal cortices increased after training with control feedback but decreased after training with punishment; however, after training on the FTT with punishment, functional connectivity between PMC with the striatum and medial frontal gyrus increased after training. Finally, we looked for regions that showed an overlap in the Rest x Feedback group interaction across tasks. At our conservative threshold, only the ventral pallidum showed an overlap for this interaction and training the two tasks reward and punishment had opposing effects on connectivity between PMC and this region. After training on the SRTT, connectivity between the PMC and the pallidum increased after training with reward but decreased after training with punishment; the opposite pattern was true after training with the FTT: connectivity between the PMC and the pallidum increased after training with punishment but decreased after training with reward. Taken together, these results suggest that training with valenced feedback has differential effects on functional connectivity after training, and that these effects primarily manifest in the regions involved in task performance.

Reward and punishment differentially impacted PMC functional connectivity with the striatum and medial temporal lobe after training on the SRTT. It is widely known that the SRTT learning engages the motor system, including motor cortex and premotor cortex (Keele SW et al. 2003; Schubotz RI and DY von Cramon 2003; Hardwick RM *et al.* 2013; Wiestler T and J Diedrichsen 2013; Kornysheva K and J Diedrichsen 2014; King BR et al. 2017), parietal cortex (Keele SW *et al.* 2003; Breton J and EM Robertson 2017), basal ganglia (Carbon M et al. 2004; Seger CA 2006; Albouy G, S Fogel, et al. 2013; Debas K et al. 2014), and medial temporal lobe (Schendan HE *et al.* 2003; Albouy G, BR King, et al. 2013). These structures are often classified as fast (frontoparietal and hippocampal) and slow (motor and basal ganglia) systems, based on the onset of evoked activity of these regions during the learning process (Doyon J *et al.* 2018), and one dominant theory to explain the impact of reward and punishment on memory formation is the engagement of the slow and fast systems, respectively. With respect to the SRTT, we found

evidence for this differential recruitment, with PMC connectivity with motor and basal ganglia structures increasing after training with reward and PMC connectivity with anterior insula and hippocampus after training with punishment. Scalp EEG recordings are consistent with this idea; during motor adaptation learning augmented with reward or punishment, beta-power, which has been associated with reward-related activity, increased in contralateral motor areas following adaptation trials where reward was delivered. In contrast, frontoparietal theta-power increased after trials where punishment was delivered (Hamel R et al. 2018). Our data suggests that this differential recruitment may propagate into the post-training period.

In addition to reward and punishment, other factors can interact with memory systems recruited during SRTT learning. For example, the contribution of the motor, parietal, and subcortical systems may be based on the goal of the learner, or based on the statistical complexity of the sequence being learned (Robertson EM 2007). However, it is clear that these memory systems interact during the learning process and that it is possible to bias activity towards one system [e.g. (Keele SW *et al.* 2003)] or to induce competitive interactions among the memory systems (Brown RM and EM Robertson 2007; Tunovic S et al. 2014) by constraining task conditions. There has been evidence to suggest that valenced feedback could also dissociate the memory systems involved during skill-learning (Wachter T *et al.* 2009), and the present data extend this knowledge by showing that reward and punishment are bias recruitment of neural areas after learning.

Despite the compelling evidence for differential engagement of fast and slow learning systems after SRTT learning, the post-training functional connectivity following training on the FTT suggests that this does not generalize across tasks. After training on the FTT, we observed differences amongst the feedback groups in functional connectivity between PMC and posterior parietal, lateral occipital, and prefrontal cortices, as well as the ventral striatum and the cerebellum. The control and to reward groups showed increases in connectivity between PMC and the parietal, occipital, and cerebellar regions, but decreasing connectivity between PMC and ventral striatal and medial frontal gyrus. Punishment showed the opposite effect:

functional connectivity between PMC and parietal, occipital, and cerebellum decreased after training and PMC connectivity with the striatum and medial frontal gyrus increased after training. FTT learning is generally associated with decreased BOLD activity in the cortical motor network, prefrontal cortex, and caudate nucleus, but increased BOLD activity in the cerebellum and putamen (Floyer-Lea A and PM Matthews 2004, 2005). While prior studies of FTT training did not find BOLD activity changes in lateral occipital cortex, this region is strongly associated with action representation (Kable JW et al. 2005; Kable JW and A Chatterjee 2006; Lingnau A and PE Downing 2015) and may be consistently engaged while learning the novel visuomotor association for learning on the FTT. In summary, Thus, the effects of feedback are largely restricted to areas involved in performance of the task.

Overall with respect to the influence of feedback on functional connectivity change due to training, we observed very little overlap between the FTT and SRTT. At our conservative threshold, the only region where feedback modulated functional connectivity change due to training after training on both tasks was the ventral pallidum. In this region, reward and punishment feedback had opposing effects in the two tasks. After training on the SRTT, functional connectivity between the PMC and ventral pallidum increased after training with reward and decreased after training with punishment. The opposite was true after training on the FTT, wherein we observed increased functional connectivity between PMC and the ventral pallidum after training with punishment, but a decrease after training with reward. The ventral pallidum contains cell populations that respond to both appetitive and aversive stimuli (Saga Y et al. 2017). Behavioral response to appetitive and aversive stimuli are mediated by glutamatergic and GABAergic neurons in the ventral pallidum, and project to the ventral tegmental area and lateral habenula, respectively (Faget L et al. 2018). Therefore, given that cells in this area respond to positive and negative stimuli, it may not surprising to find this region affected by both reward and punishment following training in our study.

An explanation for the differential activity of the pallidum across tasks may lie in the behavioral performance. Across both tasks, the dissociation between the feedback groups with respect to

PMC-pallidum functional connectivity change after training matched the dissociation behavioral performance: the group where PMC-pallium functional connectivity increased after training performed worse on the task, while the group performing better at the end of training showed diminished PMC-pallidum functional connectivity [see (Steel A *et al.* 2016) for behavioral results of the participants in this study]. Specifically, on the SRTT, the punishment group performed better during training (i.e. had a faster overall reaction time) compared to the reward and control groups, and, on this task we observed a decrease in functional connectivity between ventral pallidum and PMC after training with punishment. On the other hand, in the FTT, the punishment group performed worse (i.e. made more tracking error) than the reward and control groups behaviorally, and, in this case, the connectivity between PMC and ventral pallidum increased after training with punishment. Thus, with respect to the pallidum, it may be that the connectivity of this region to PMC depends on the experience of the learner (either absence of reward or presence of punishment) in the environment, and this activity may persist after training. In line with this hypothesis, a meta-analysis of decision making studies using valenced feedback found surprise, valence, and signed prediction error information converged specifically in the ventral pallidum (Fouragnan E *et al.* 2018).

Even at a liberal threshold, the regions that showed overlap across the two tasks showed opposite patterns of connectivity change based on the feedback given. Similar to the pallidum, the differential effect of feedback on functional connectivity between PMC and these regions may be due to the qualitative differences in performance. More generally, the minimal overlap and opposing patterns suggest that reward and punishment have strong impacts on post-training brain activity and that this effect is strongly influenced by the task that was performed.

Valenced feedback processing is often characterized as being mediated by a set of regions that would be common across tasks; for example, in the context of statistical learning and decision making, distinct networks for valence, surprise, and signed-prediction error have been suggested (Fouragnan E *et al.* 2018). Regions activated in response to surprise and valence information include the striatum and insular cortex, which differentially responded to positive

and negative surprises, respectively (Fouragnan E *et al.* 2018). The regions showing changes in functional connectivity after training on the SRTT task match these predictions, and, indeed, the SRTT may be considered a statistical learning task (Robertson EM 2007). However, after training on the FTT, the patterns of connectivity change across the feedback groups was not consistent with what would be expected after statistical learning. After training on FTT, we observed that task-relevant regions, including the lateral occipital, parietal, and cerebellar activity (Imamizu H et al. 2000; Krakauer JW et al. 2004; Grafton ST et al. 2008; Hardwick RM *et al.* 2013) showed feedback related changes in functional connectivity. The absence of basal gangliar modulation in this task, given the role of the basal ganglia in motivation-based adjustments of movement and performance. However, there is some evidence that the cerebellum also plays a role in motivational aspects of movement correction during motor learning (Turner RS et al. 2003; Turner RS and M Desmurget 2010; Bostan AC and PL Strick 2018), and it is likely that this depends on task-demands, as well. In summary, our data provide further evidence that, rather than utilize a core feedback-processing network segregated on the basis of valence, rewards and punishments may be processed in a distributed fashion amongst brain regions engaged by the task.

Several aspects of the present study are worth noting. First, we did not distinguish between dorsal and ventral premotor cortex in our analyses because no regions showed a Rest x Group x ROI interaction in either task. Dorsal and ventral premotor cortex are highly interconnected and although feedback may differentially impact dorsal and ventral premotor cortex, we may have been underpowered to detect any effect. Second, it should be noted that our implementation of the SRTT included a fixed trial length rather than a self-paced trial duration, which might foster explicit knowledge, and it has been shown that explicit learning recruits different neural networks after learning (Sami S et al. 2014). No participants included in the study spontaneously reported sequence knowledge. When tested at 3+ weeks, participants showed no evidence of explicit awareness (for further discussion, see Steel et al., 2016a). In addition, in order to match the SRTT parameters, our force tracking task implementation featured continuous feedback of cursor position, rather than feedback at given at the end of the trial,

and future work may consider the difference between continuous versus end-of-trial feedback. Third, we did not find any differences in retention across the feedback groups in either task, despite observing differences in the acquisition and performance during training. Therefore, we are not able to say whether any specific PMC connections facilitate better memory formation. Finally, it is widely known that sleep interacts with memory formation and offline memory processes (Robertson EM, A Pascual-Leone and RC Miall 2004; Robertson EM, A Pascual-Leone and DZ Press 2004; Albouy G, S Fogel*, et al.* 2013). In our study, we did not measure sleep quality. Whether sleep interacts with the effect of reward and punishment on retention may be an interesting avenue for future research.

In summary, we found that reward and punishment have distinct effects on changes in resting state functional connectivity due to training. This suggests that rather than a critical feedback-processing network, feedback may be processed within those systems specifically engaged by the task requirements.

References

Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG. 2011. Reward Improves Long-Term Retention of a Motor Memory through Induction of Offline Memory Gains. Curr Biol 21:557-562.

Albouy G, Fogel S, Pottiez H, Nguyen VA, Ray L, Lungu O, Carrier J, Robertson E, Doyon J. 2013. Daytime sleep enhances consolidation of the spatial but not motoric representation of motor sequence memory. PloS one 8:e52805.

Albouy G, King BR, Maquet P, Doyon J. 2013. Hippocampus and striatum: dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. Hippocampus 23:985-1004.

Bostan AC, Strick PL. 2018. The basal ganglia and the cerebellum: nodes in an integrated network. Nature reviews Neuroscience.

Breton J, Robertson EM. 2017. Dual enhancement mechanisms for overnight motor memory consolidation. Nat Hum Behav 1.

Brown RM, Robertson EM. 2007. Off-line processing: reciprocal interactions between declarative and procedural memories. The Journal of neuroscience : the official journal of the Society for Neuroscience 27:10468-10475.

Carbon M, Ma Y, Barnes A, Dhawan V, Chaly T, Ghilardi MF, Eidelberg D. 2004. Caudate nucleus: influence of dopaminergic input on sequence learning and brain activation in Parkinsonism. NeuroImage 21:1497-1507.

Chen G, Saad ZS, Britton JC, Pine DS, Cox RW. 2013. Linear mixed-effects modeling approach to FMRI group analysis. NeuroImage 73:176-190.

Cox RW. 1996. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162-173.

Debas K, Carrier J, Barakat M, Marrelec G, Bellec P, Hadj Tahar A, Karni A, Ungerleider LG, Benali H, Doyon J. 2014. Off-line consolidation of motor sequence learning results in greater integration within a cortico-striatal functional network. NeuroImage 99:50-58.

Doyon J, Gabitov E, Vahdat S, Lungu O, Boutin A. 2018. Current issues related to motor sequence learning in humans. Current Opinion in Behavioral Sciences 20:89-97.

Evans JW, Kundu P, Horovitz SG, Bandettini PA. 2015. Separating slow BOLD from non-BOLD baseline drifts using multi-echo fMRI. NeuroImage 105:189-197.

Faget L, Zell V, Souter E, McPherson A, Ressler R, Gutierrez-Reed N, Yoo JH, Dulcis D, Hnasko TS. 2018. Opponent control of behavioral reinforcement by inhibitory and excitatory projections from the ventral pallidum. Nature communications 9:849.

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33:341-355.

Floyer-Lea A, Matthews PM. 2004. Changing brain networks for visuomotor control with increased movement automaticity. Journal of neurophysiology 92:2405-2412.

Floyer-Lea A, Matthews PM. 2005. Distinguishable brain activation networks for short- and long-term motor skill learning. Journal of neurophysiology 94:512-518.

Floyer-Lea A, Wylezinska M, Kincses T, Matthews PM. 2006. Rapid modulation of GABA concentration in human sensorimotor cortex during motor learning. Journal of neurophysiology 95:1639-1644.

Fouragnan E, Retzler C, Philiastides MG. 2018. Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. Hum Brain Mapp.

Galea JM, Mallia E, Rothwell J, Diedrichsen J. 2015. The dissociable effects of punishment and reward on motor learning. Nature neuroscience 18:597-602.

Gonzalez-Castillo J, Handwerker DA, Robinson ME, Hoy CW, Buchanan LC, Saad ZS, Bandettini PA. 2014. The spatial structure of resting state connectivity stability on the scale of minutes. Frontiers in neuroscience 8:138.

Gotts SJ, Simmons WK, Milbury LA, Wallace GL, Cox RW, Martin A. 2012. Fractionation of social brain circuits in autism spectrum disorders. Brain : a journal of neurology 135:2711-2725.

Grafton ST, Schmitt P, Van Horn J, Diedrichsen J. 2008. Neural substrates of visuomotor learning based on improved feedback control and prediction. NeuroImage 39:1383-1395.

Gruber MJ, Ritchey M, Wang SF, Doss MK, Ranganath C. 2016. Post-learning Hippocampal Dynamics Promote Preferential Retention of Rewarding Events. Neuron 89:1110-1120.

Hamel R, Savoie FA, Lacroix A, Whittingstall K, Trempe M, Bernier PM. 2018. Added value of money on motor performance feedback: Increased left central beta-band power for rewards and fronto-central theta-band power for punishments. NeuroImage 179:63-78.

Hardwick RM, Lesage E, Eickhoff CR, Clos M, Fox P, Eickhoff SB. 2015. Multimodal connectivity of motor learning-related dorsal premotor cortex. NeuroImage 123:114-128.

Hardwick RM, Rottschy C, Miall RC, Eickhoff SB. 2013. A quantitative meta-analysis and review of motor learning in the human brain. NeuroImage 67:283-297.

Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Putz B, Yoshioka T, Kawato M. 2000. Human cerebellar activity reflecting an acquired internal model of a new tool. Nature 403:192-195.

Jo HJ, Gotts SJ, Reynolds RC, Bandettini PA, Martin A, Cox RW, Saad ZS. 2013. Effective Preprocessing Procedures Virtually Eliminate Distance-Dependent Motion Artifacts in Resting State FMRI. Journal of applied mathematics 2013.

Jo HJ, Saad ZS, Simmons WK, Milbury LA, Cox RW. 2010. Mapping sources of correlation in resting state FMRI, with artifact detection and removal. NeuroImage 52:571-582.

Kable JW, Chatterjee A. 2006. Specificity of action representations in the lateral occipitotemporal cortex. Journal of cognitive neuroscience 18:1498-1517.

Kable JW, Kan IP, Wilson A, Thompson-Schill SL, Chatterjee A. 2005. Conceptual representations of action in the lateral temporal cortex. Journal of cognitive neuroscience 17:1855-1870.

Keele SW, Ivry R, Mayr U, Hazeltine E, Heuer H. 2003. The cognitive and neural architecture of sequence representation. Psychol Rev 110:316-339.

King BR, Saucier P, Albouy G, Fogel SM, Rumpf JJ, Klann J, Buccino G, Binkofski F, Classen J, Karni A, Doyon J. 2017. Cerebral Activation During Initial Motor Learning Forecasts Subsequent Sleep-Facilitated Memory Consolidation in Older Adults. Cerebral cortex 27:1588-1601.

Kornysheva K, Diedrichsen J. 2014. Human premotor areas parse sequences into their spatial and temporal features. eLife 3:e03043.

Krakauer JW, Ghilardi MF, Mentis M, Barnes A, Veytsman M, Eidelberg D, Ghez C. 2004. Differential cortical and subcortical activations in learning rotations and gains for reaching: a PET study. Journal of neurophysiology 91:924-933.

Kundu P, Santin MD, Bandettini PA, Bullmore ET, Petiet A. 2014. Differentiating BOLD and non-BOLD signals in fMRI time series from anesthetized rats using multi-echo EPI at 11.7 T. NeuroImage 102 Pt 2:861-874.

Lingnau A, Downing PE. 2015. The lateral occipitotemporal cortex in action. Trends Cogn Sci 19:268-277.

Murty VP, Labar KS, Adcock RA. 2012. Threat of punishment motivates memory encoding via amygdala, not midbrain, interactions with the medial temporal lobe. The Journal of neuroscience : the official journal of the Society for Neuroscience 32:8969-8976.

Murty VP, LaBar KS, Adcock RA. 2016. Distinct medial temporal networks encode surprise during motivation by reward versus punishment. Neurobiology of learning and memory 134 Pt A:55-64.

Murty VP, Tompary A, Adcock RA, Davachi L. 2016. Selectivity in post-encoding connectivity with high-level visual cortex is associated with reward-motivated memory. The Journal of neuroscience : the official journal of the Society for Neuroscience.

Mushiake H, Inase M, Tanji J. 1991. Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. Journal of neurophysiology 66:705-718.

Peterson EJ, Seger CA. 2013. Many hats: intratrial and reward level-dependent BOLD activity in the striatum and premotor cortex. Journal of neurophysiology 110:1689-1702.

Poser BA, Versluis MJ, Hoogduin JM, Norris DG. 2006. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI. Magnetic resonance in medicine 55:1227-1235.

Posse S, Wiese S, Gembris D, Mathiak K, Kessler C, Grosse-Ruyken ML, Elghahwagi B, Richards T, Dager SR, Kiselev VG. 1999. Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. Magnetic resonance in medicine 42:87-97.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59:2142-2154.

Quattrocchi G, Greenwood R, Rothwell JC, Galea JM, Bestmann S. 2017. Reward and punishment enhance motor adaptation in stroke. Journal of neurology, neurosurgery, and psychiatry.

Ramkumar P, Dekleva B, Cooler S, Miller L, Kording K. 2016. Premotor and Motor Cortices Encode Reward. PloS one 11:e0160851.

Robertson EM. 2007. The serial reaction time task: implicit motor skill learning? The Journal of neuroscience : the official journal of the Society for Neuroscience 27:10073-10075.

Robertson EM, Pascual-Leone A, Miall RC. 2004. Current concepts in procedural consolidation. Nature reviews Neuroscience 5:576-582.

Robertson EM, Pascual-Leone A, Press DZ. 2004. Awareness modifies the skill-learning benefits of sleep. Curr Biol 14:208-212.

Saad ZS, Reynolds RC, Jo HJ, Gotts SJ, Chen G, Martin A, Cox RW. 2013. Correcting brain-wide correlation differences in resting-state FMRI. Brain connectivity 3:339-352.

Saga Y, Richard A, Sgambato-Faure V, Hoshi E, Tobler PN, Tremblay L. 2017. Ventral Pallidum Encodes Contextual Information and Controls Aversive Behaviors. Cerebral cortex 27:2528-2543.

Sami S, Robertson EM, Miall RC. 2014. The time course of task-specific memory consolidation effects in resting state networks. The Journal of neuroscience : the official journal of the Society for Neuroscience 34:3982-3992.

Schendan HE, Searl MM, Melrose RJ, Stern CE. 2003. An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron 37:1013-1025.

Schubotz RI, von Cramon DY. 2003. Functional-anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. NeuroImage 20:S120-S131.

Seger CA. 2006. The basal ganglia in human learning. Neuroscientist 12:285-290.

Shigemune Y, Tsukiura T, Kambara T, Kawashima R. 2014. Remembering with gains and losses: effects of monetary reward and punishment on successful encoding activation of source memories. Cerebral cortex 24:1319-1331.

Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW. 2012. Overcoming motor "forgetting" through reinforcement of learned actions. The Journal of neuroscience : the official journal of the Society for Neuroscience 32:14617-14621.

Steel A, Silson EH, Stagg CJ, Baker CI. 2016. The impact of reward and punishment on skill learning depends on task demands. Scientific reports 6:36056.

Therrien AS, Wolpert DM, Bastian AJ. 2016. Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. Brain : a journal of neurology 139:101-114.

Tomassini V, Jbabdi S, Klein JC, Behrens TE, Pozzilli C, Matthews PM, Rushworth MF, Johansen-Berg H. 2007. Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. The Journal of neuroscience : the official journal of the Society for Neuroscience 27:10259-10269.

Tunovic S, Press DZ, Robertson EM. 2014. A physiological signal that prevents motor skill improvements during consolidation. The Journal of neuroscience : the official journal of the Society for Neuroscience 34:5302-5310.

Turner RS, Desmurget M. 2010. Basal ganglia contributions to motor control: a vigorous tutor. Current opinion in neurobiology 20:704-716.

Turner RS, Desmurget M, Grethe J, Crutcher MD, Grafton ST. 2003. Motor subcircuits mediating the control of movement extent and speed. J Neurophysiol 90:3958-3966.

Wachter T, Lungu OV, Liu T, Willingham DT, Ashe J. 2009. Differential effect of reward and punishment on procedural learning. The Journal of neuroscience : the official journal of the Society for Neuroscience 29:436-443.

Wiestler T, Diedrichsen J. 2013. Skill learning strengthens cortical representations of motor sequences. eLife 2:e00801.

Tables

Number of voxels	Volume	x	y	z	Region
SRTT					
Main effect Rest					
1909	51543	-4.5	-22.5	11.5	Left Thalamus
949	25623	52.5	1.5	-6.5	Right superior temporal gyrus
477	12879	1.5	-73.5	-24.5	Right cerebellar vermis
299	8073	-55.5	-1.5	-3.5	Left superior temporal gyrus
68	1836	25.5	-13.5	-12.5	Right parahippocampal cortex
Rest x Condition					
828	22356	-16.5	1.5	-0.5	Left pallidum
215	5805	28.5	1.5	-18.5	Right parahippocampal cortex
195	5265	-10.5	-34.5	59.5	Left supplementary motor area
154	4158	-4.5	-37.5	-24.5	Left cerebelllum
77	2079	-16.5	-13.5	-18.5	Left parahippocampal cortex
59	1593	-31.5	19.5	15.5	Left inferior frontal gyrus
FTT					
Main effect Rest					
3545	95715	-4.5	-22.5	11.5	Left thalamus
148	3996	7.5	-31.5	59.5	Right supplementary motor area
118	3186	-1.5	55.5	-6.5	Left medial frontal gyrus
112	3024	43.5	-61.5	-3.5	Left middle occipital gyrus
112	3024	1.5	-61.5	17.5	Right posterior cingulate cortex
60	1620	16.5	-73.5	29.5	Right precuneus
Rest x Condition					
639	17253	46.5	-49.5	-15.5	Right fusiform gyrus
395	10665	-49.5	-64.5	8.5	Left middle temporal gyrus
147	3969	-10.5	-91.5	5.5	Left cuneus
124	3348	25.5	-73.5	38.5	Right precuneus
74	1998	13.5	1.5	5.5	Right pallidum
63	1701	19.5	-61.5	-21.5	Right cerebellum lobule VI

60	1620	-7.5	-64.5	2.5	Left lingual gyrus
55	1485	7.5	61.5	2.5	Right medial frontal gyrus

Figures

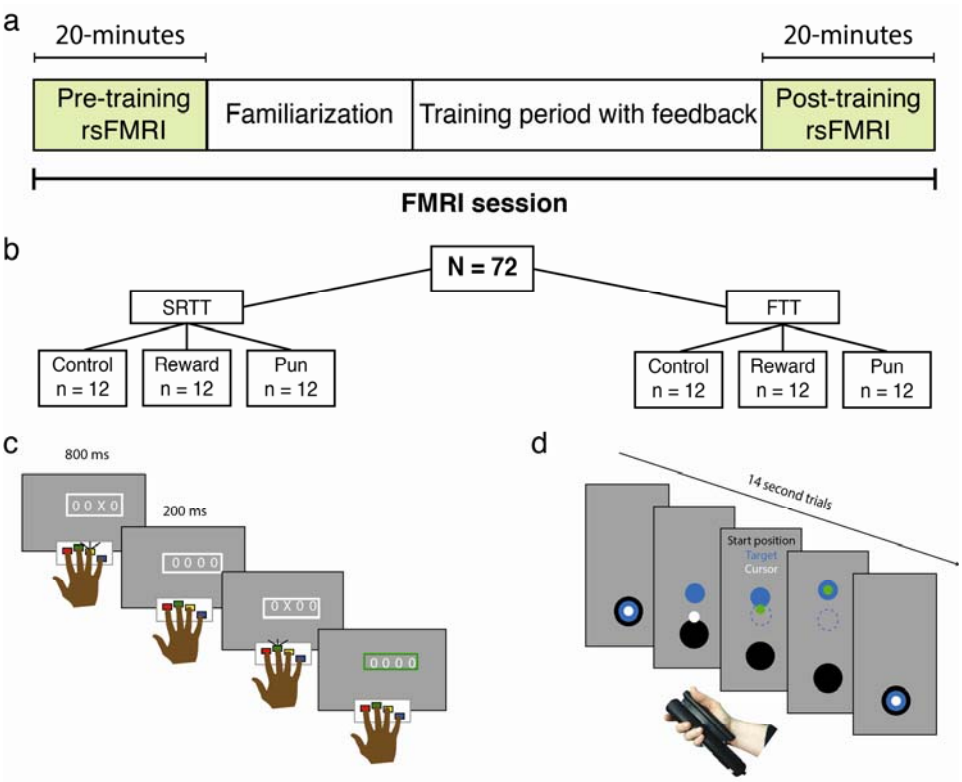

Figure 1. Experimental design and skill memory retention. (a,b) Participants underwent 20 minutes of resting state fMRI before and after training on either the serial reaction time task (SRTT) or the force tracking task (FTT) while receiving REW, PUN, or CONT feedback. In the SRTT (c), participants responded to a cue appearing in one of four locations on a screen. In the FTT (d), participants modulated their grip force to track a moving target. In both tasks, the stimulus could follow either a random or fixed sequence, and skill memory was assessed by comparing performance during random and fixed trials.

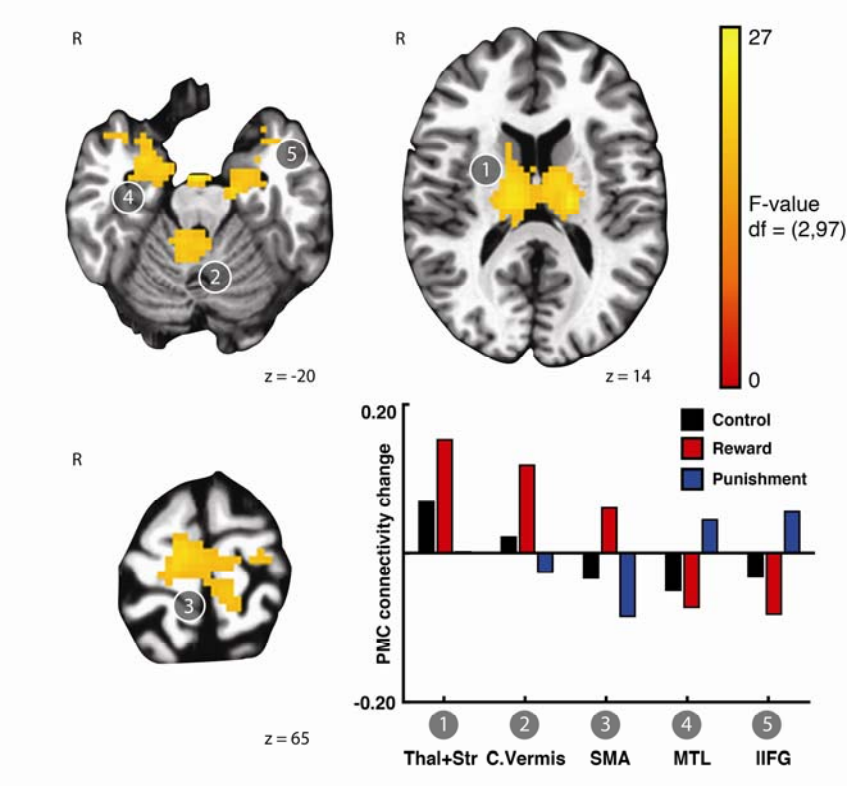

Figure 2. Reward and punishment differentially effect PMC functional connectivity change after training on the SRTT. Linear mixed effects model revealed brain regions exhibiting a Rest x Feedback group interaction in the functional connectivity of left PMC. Functional connectivity between PMC and the thalamus and striatum, cerebellum, and SMA increased after training with reward and control feedback but decreased after training with punishment. In contrast, functional connectivity between medial temporal lobe and left inferior frontal gyrus increased after training with punishment but decreased after training with reward and control feedback. Bars show mean connectivity change after training estimated by the linear mixed effects model.

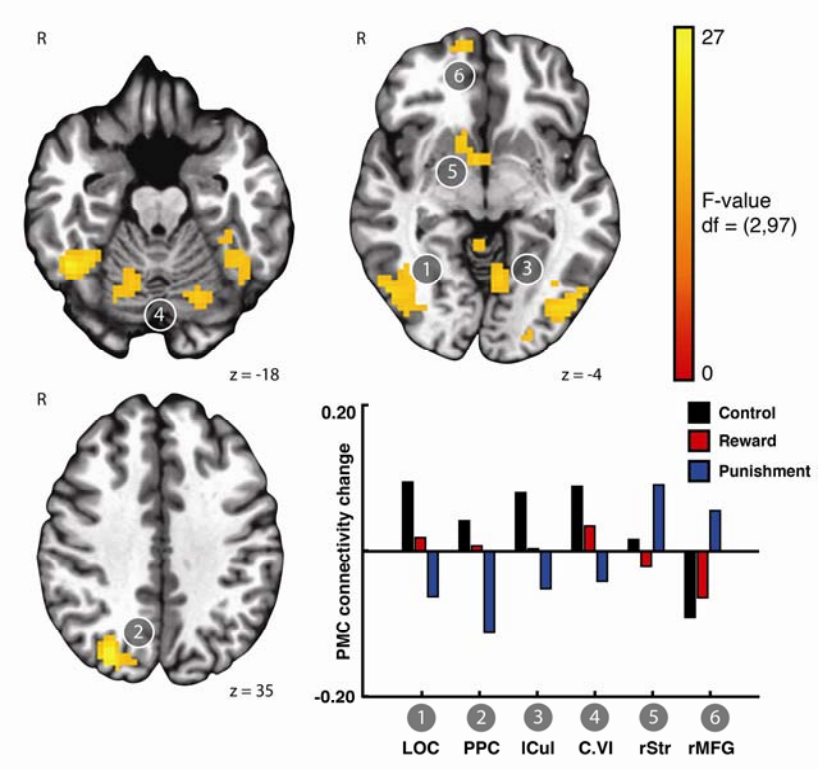

Figure 3. Punishment promotes PMC-striatal connectivity after FTT training. Linear mixed effects model revealed brain regions exhibiting a Rest x Feedback group interaction in the functional connectivity of left PMC. Functional connectivity between the PMC and lateral occipital cortex, posterior parietal cortex, culmen, and cerebellum increased after training with control feedback, but decreased after training with punishment. In contrast, PMC connectivity with Right ventral striatum and right medial frontal gyrus increased after training with punishment but decreased after training with reward and control feedback. Bars show mean connectivity change after training estimated by the linear mixed effects model.

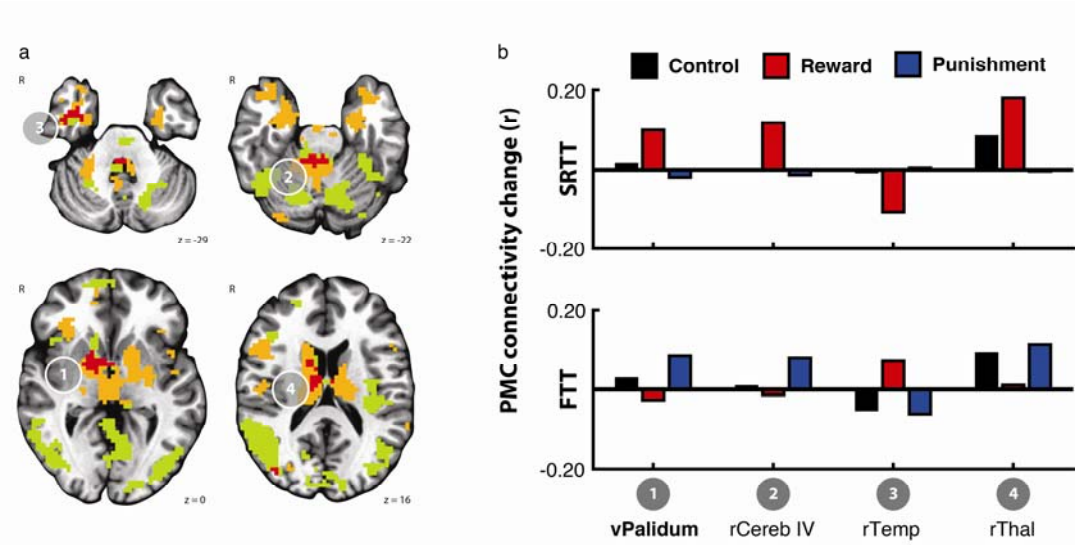

Figure 4. Regions showing PMC functional connectivity change after training were had limited overlap across tasks. Clusters that showed a PMC connectivity change after training that differed by the feedback group in the SRTT (orange) and the FTT (green) showed limited overlap, even at a liberal cluster threshold ($p < 0.01$, $k = 54$). Only the ventral pallidum (Cluster 1, bold) showed an overlap at the threshold reported above ($p < 0.005$, $k = 54$). In all clusters, the Reward and Punishment groups exhibited opposite patterns of connectivity change with PMC due to training on the SRTT (upper) and FTT (lower). Bars show mean connectivity change after training estimated by the linear mixed effects model.

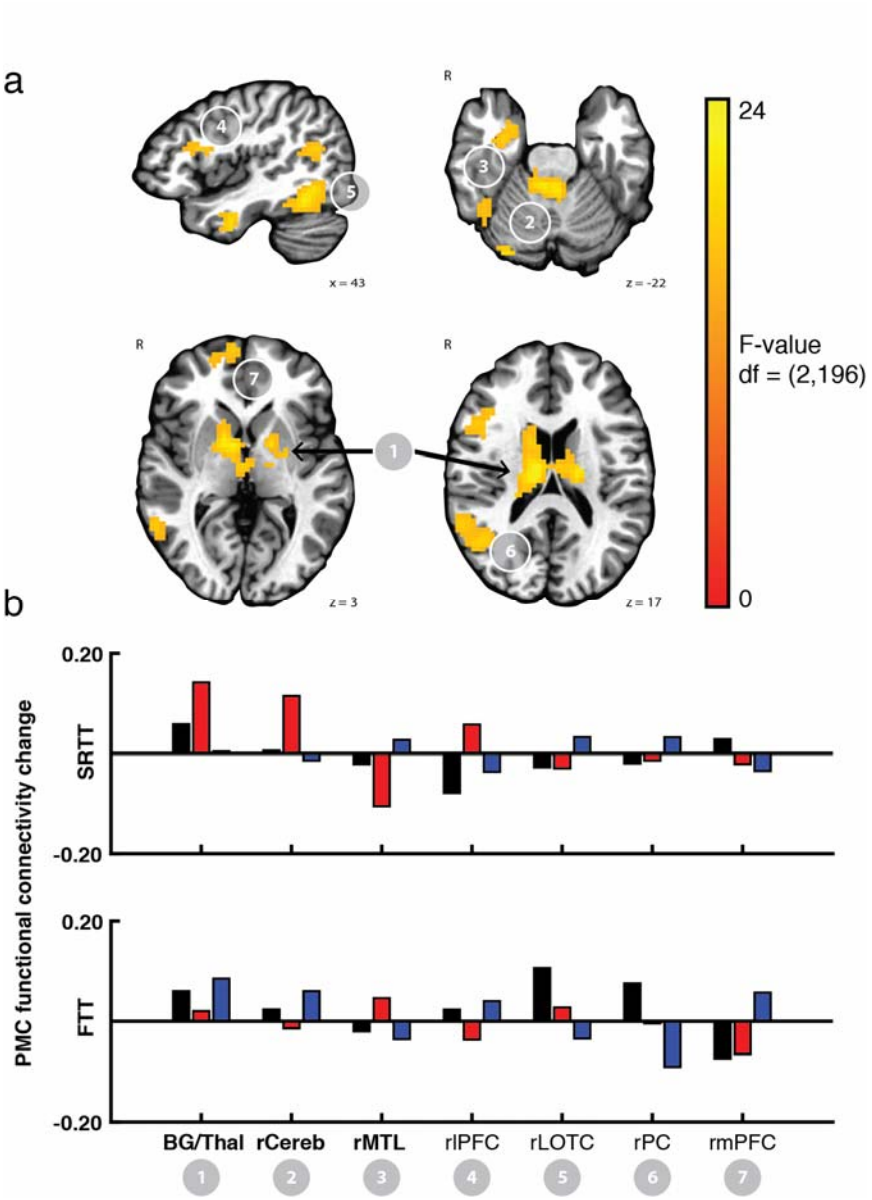

Figure S1. Difference in functional connectivity change after training with feedback depends on task. Upper. Data from both tasks were submitted to a linear mixed effects model. This analysis revealed the resulting F-statistic map for the interaction term 'Task*Group*Rest' indicating regions within the basal ganglia, cerebellum, MTL, temporal cortex, parietal cortex, and prefrontal cortex showed an effect of feedback on functional connectivity change after training that varied by task. Data from these ROIs is shown in the lower panel. Importantly, the ROIs that showed an overlap in Figure 4 evidenced the interaction shown here.