bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Reward and punishment differentially recruit cerebellum and medial temporal lobe to facilitate skill memory retention

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: 200

Introduction word count: 521

Discussion word count: 1430

Number of pages: 39

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

Acknowledgements: The authors would like to thank Matthew Rushworth for his helpful comments on the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Abstract

Reward and punishment shape behavior, but the neural mechanisms underlying their effect on skill learning are not well understood. The premotor cortex (PMC) is known to play a central role in sequence learning and has a diverse set of structural and connections with cortical (e.g. medial temporal/parietal lobes) and subcortical (caudate/cerebellum) memory systems that might be modulated by valenced feedback. Here, we tested whether the functional connectivity of PMC immediately after training with reward or punishment predicted memory retention across two different tasks. Resting-state fMRI was collected before and after 72 participants trained on either a serial reaction time or force-tracking task with reward, punishment, or control feedback. Training-related change in PMC functional connectivity was compared across feedback groups. Reward and punishment differentially affected PMC functional connectivity: PMC-cerebellum connectivity increased following training with reward, while PMC-medial temporal lobe connectivity increased after training with punishment. Moreover, feedback impacted the relationship between PMC-caudate connectivity and 24-48hour skill memory. These results were consistent across the tasks, suggestive of a general, non-task-specific mechanism by which feedback modulates skill learning. These findings illustrate dissociable roles for the medial temporal lobe and cerebellum in skill memory retention and suggest novel ways to optimize behavioral training.

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

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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, EH Silson, et al. 2016). Prior behavioral studies of motor adaptation, suggest differences between reward and punishment. 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 also restored adaptation learning in patients with cerebellar degeneration (Therrien AS et al. 2016) and stroke (Quattrocchi G et al. 2017). However, results in non-cerebellar-dependent tasks have been mixed, with reward either increasing offline gains (Abe M *et al.* 2011) or having no effect (Steel A, EH Silson, *et al.* 2016).

One parsimonious explanation for the conflicting findings about the effects of valenced feedback on non-cerebellar-dependent tasks may be 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). Others have shown memory benefits for reward mediated by MTL in episodic memory, but this has not been replicated for sequence learning (Gruber MJ et al. 2016; Murty VP, A Tompary, et al. 2016). However, these studies only focused on brain activity *during* performance. For valenced feedback to be effective in modulation skill retention, it is important to understand its effects on the off-line mechanisms that facilitate long-term memory storage.

Thus, we sought to determine how valenced feedback affects neural processing immediately *after* training. Participants trained on one of two skill learning tasks

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

(serial reaction time task [SRTT] or force tracking task [FTT]) augmented with reward, punishment, or uninformative (control) feedback. Before and after training we collected 20-minutes of resting-state fMRI data (Figure 1a-d;). Outside of the scanner, we tested participants' recall of the learned sequence at 1-hour (1-h), 24-48 hours (24-48h), and 3+ weeks after training.

In the behavioral data, which was published previously (Steel A, EH Silson, *et al.* 2016), we found strong differences in skill acquisition between feedback conditions but no effect of feedback on retention. The disparity between acquisition and retention suggests different mechanisms may stabilize memory across the feedback conditions. The premotor cortex (PMC) is a critical memory-encoding region for sequence learning (Hardwick RM et al. 2013), and shows reward-related activity after movement (Ramkumar P et al. 2016). Further, given its connections to motor, parietal, and prefrontal cortices (Tomassini V et al. 2007), there are multiple ways in which feedback could modulate processing within the PMC. We therefore 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.

Materials and methods

Overview

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, EH Silson, *et al.* 2016) and is summarized below. Training on each task was conducted while participants underwent fMRI scanning. Before and after the training session, 20-minutes of resting-state fMRI was collected. To investigate retention of skill memory, subjects were tested 1-h, 24-48h, and 3+ weeks after the initial training. We were primarily

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

interested in the effects of feedback on memory retention at 24-48h. We chose the 24-48h rather than the 3+ weeks probe, as the latter will be influenced by long-term memory decay processes unrelated to the period immediately following learning when the resting state acquisition occurred. When examining the relationship between functional connectivity change and performance at 24-48h, we included the 1-h and 3+ weeks memory tests as covariates in our imaging model, ensuring that any effects we observed at 24-48h were specific to this time period. The 24-48h probe always occurred after at least one night's sleep.

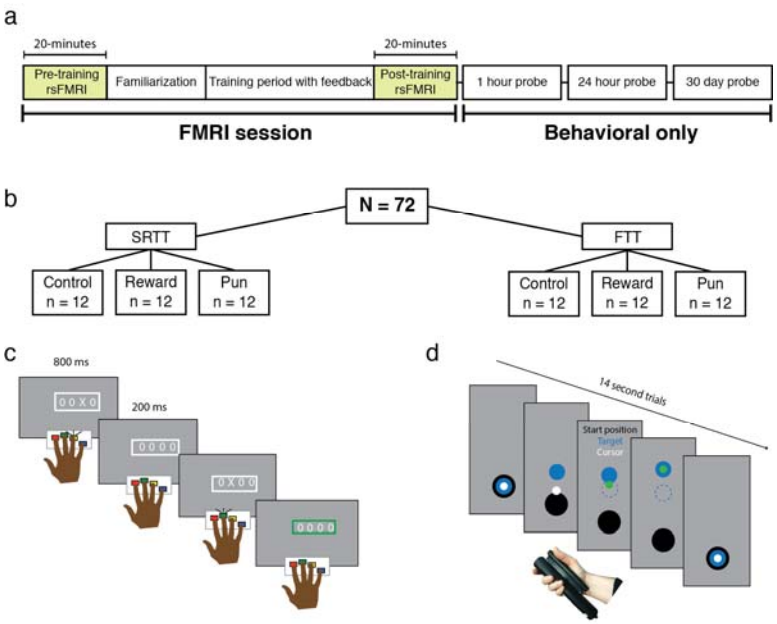

Figure 1. Experimental design and skill memory retention. (a,b) Participants underwent 20 minutes of resting state fMRI before and after training on either the serial reaction time task (SRTT) or the force tracking task (FTT) while receiving REW, PUN, or CONT feedback. Participants were then tested for skill memory 1-h, 24-48h, and 3+ weeks after initial training. 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.

Participants

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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.

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 SRT 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.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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.

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

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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. Between each block, participants were shown their cumulative earnings. Control participants saw the phrase "You have money."

Behavioral statistical analysis

The present study deals with the relationship of the pre- and post- training resting state brain activity to the retention of the skill memory at 1-h, 24-48h, and 3+ weeks. For a detailed description of the behavioral data collected during training and retention, see Steel A, EH Silson, *et al.* (2016). Here, given our primary

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

hypothesis concerning the mechanisms underpinning skill retention, we focused on the retention data only. The behavioral performance measures in the two tasks differed in scale, as the SRT task focuses on RT, while the FTT is based on distance from the target. Here, in order to pool the data across both tasks, participants were ranked based on their relative performance. Rank normalization is common practice in non-parametric statistics, for example Spearman's rho. For each task separately, the 36 unique participants were ranked based on their sequence knowledge at each delayed test probe. Thus, in each task and for each time point, the worst participant was ranked 1 and the best was ranked 36. These ranks were used as behavioral covariates. Using ranks rather than raw behavioral measures did not alter our previously reported behavioral results.

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

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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 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:

$$\mathrm{S}(\mathrm{TE_n}) = \mathrm{S_0} exp(-\mathrm{R_2} * \mathrm{TE_n}),$$

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.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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

We used multiple approaches to the analysis of the resting state MRI data: 1) focused on anatomically defined regions of interest and 2) model-free voxel-wise analysis (global connectedness (Gotts SJ *et al.* 2012; Song S et al. 2015; Steel A, S Song, et al. 2016)).

For all group tests, the average smoothness of the data was estimated (3dFWHMx). Data were corrected for multiple comparisons using Monte-Carlo simulations to

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

(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

Left dorsal and ventral premotor cortex (PMd and PMv) were defined based on a publically available diffusion-MRI based parcellation of premotor cortex (Tomassini V *et al.* 2007). These ROIs were chosen based on their importance in motor learning and motor control in both sensorimotor learning and sequence learning (Hardwick RM *et al.* 2013; Hardwick RM et al. 2015). We focused on the left PMC because the participants were performing the task with their right hand. 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. The resulting maps were then submitted to a linear mixed effects model (3dLME) with ROI (PMd/PMv), Rest (pre-/post-), Group (Control/Reward/Punishment), and Task (SRT/FTT).

To determine whether the relationship of the brain regions discovered above (i.e. those significant regions for the model term Rest x Group, with behavior differed between the feedback groups, we performed a partial correlation of the behavioral performance of each participant with the connectivity values from the regions discovered above, controlling for movement, global correlation, as well as 1-h and 3+ weeks behavior. We compared the correlation values across Feedback Groups, correcting for multiple comparisons. We also performed a whole brain analysis that included the factors listed above and behavioral covariates (performance during the 1-h, 24-48h, and 3+ weeks tests) to determine whether any regions showed significantly different correlation with behavior across groups across the brain. Global correlation (Saad ZS et al. 2013) (@compute_gcor) and magnitude of motion across runs (@1dDiffMag) were included as nuisance covariates for this test.

Caudate functional connectivity

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Because of the involvement of the caudate in early sequence learning and feedback processing (Carbon M et al. 2004; Seger CA 2006; Stillman CM et al. 2013), we performed a ROI analysis of the connectivity between premotor cortex and the caudate. The left caudate was defined anatomically using the TT-N27 atlas Eickhoff-Zilles maximum probability maps (Eickhoff SB et al. 2006) included in the AFNI software package. The mean connectivity between this caudate mask and left premotor ventral and dorsal was calculated on a participant-by-participant basis and compared using a linear mixed effects model with Rest (pre-/post- training), Group (Control/Reward/Punishment), and Task (SRT/FTT) as factors. To examine the relationship between this region and behavior, we calculated the correlation between the connectivity change due to training and behavioral performance at 24-48h, while controlling for motion and global correlation.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Results

To identify task-independent effects of feedback on functional connectivity immediately following training, we used a seed-based analysis focused on the left PMC. Firstly, we identified regions where the change in functional connectivity with the PMC after training differed across the feedback groups. Secondly, we then investigated the behavioural relevance of these changes, by examining whether changes in functional connectivity with PMC after training predicted 24-48h memory retention.

To address both these questions, we first implemented a voxel-wise linear mixed effects (LME) model (Chen G et al. 2013) with Rest (Pre-/Post- training), Group (Reward [REW]/Punishment [PUN]/Control [CONT]), Task (SRTT/FTT) and ROI (ventral-/dorsal- premotor cortex) as factors. This analysis revealed no differences between dorsal and ventral PMC, so we collectively refer to them as 'PMC'. Likewise, there was no interaction between Group and Task, indicating the same pattern of results across both tasks and thus the data from both tasks are considered together. A full description of the results from this model is available in Table 1.

Reward and punishment engage dissociable networks after learning.

After training, all groups showed enhanced connectivity between PMC and subcortical structures including thalamus and basal ganglia (LME: Main effect of Rest). However, given we were primarily interested in the effect of feedback on the connectivity change induced by training, we focused on the interaction between rest period (pre- versus post-) and feedback (LME: Rest x Group interaction). Seven regions exhibited a significant interaction (Figure 2, upper) with three different profiles: (i) parietal and bilateral occipitotemporal cortices became more functionally connected with left PMC after training with CONT compared to REW and PUN; (ii) Supplementary motor area and cerebellum became more connected to left PMC after training with REW, and (iii) anterior insula, immediately adjacent to inferior frontal gyrus, and anterior MTL became more connected to left PMC after training with PUN (Figure 2, lower; Figure 3 for description of the MTL region).

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

To ensure this result was not solely dependent on our *a priori* anatomical ROIs, we repeated the analysis using a functionally defined motor ROI (based on activity during the training period) that encompassed bilateral motor cortex, premotor cortex, and SMA. Using this functionally-defined ROI, we found that functional connectivity increased with occipitotemporal cortex after training with CONT feedback, the putamen and cerebellum became more connected to the functional ROI after training with REW, and connectivity between the motor ROI and anterior insula, prefrontal cortex, and MTL increased after training with PUN, consistent with our anatomical PMC ROI results.

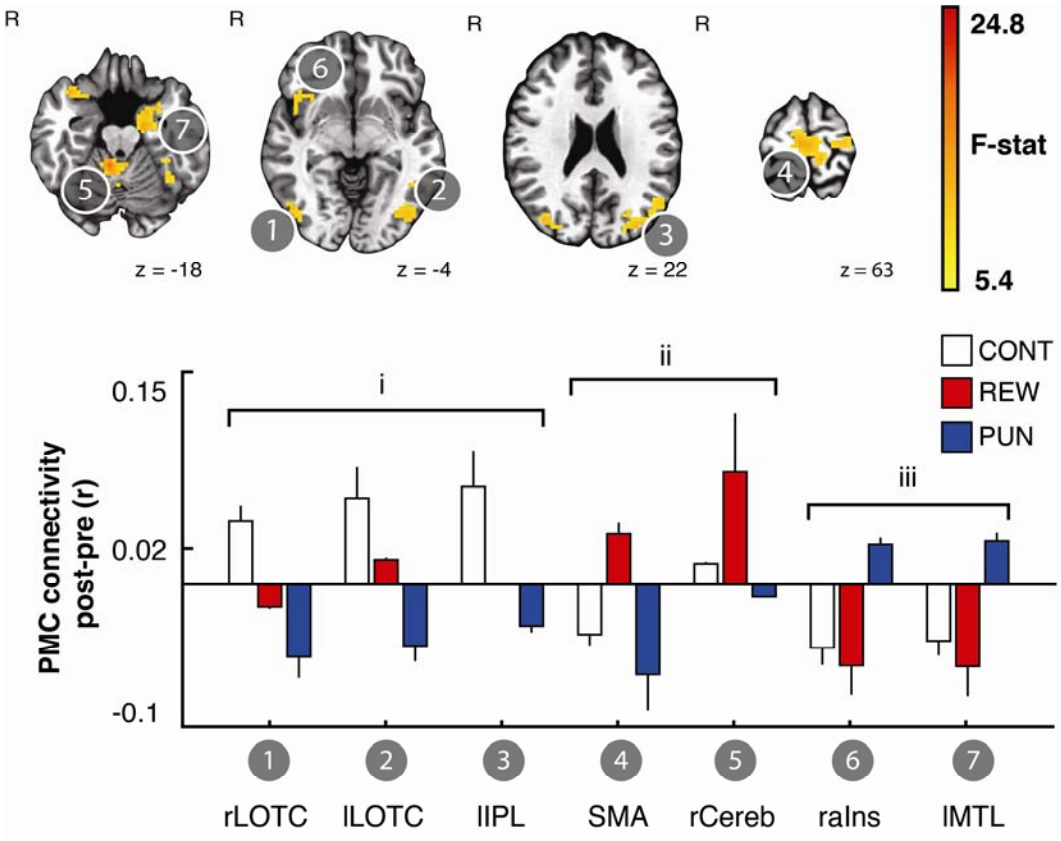

Figure 2. Premotor cortices connectivity change varies by feedback group. (*Right*) Linear mixed effects model revealed brain regions exhibiting a Rest x Feedback group interaction. (Left) These included parietoccipital regions (i), where connectivity increased after training with CONT feedback, supplementary motor area and cerebellum (ii), where connectivity increased after training with REW, and MTL and anterior insula (iii), where connectivity increased after training with PUN. Bar chart shows

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

mean ± SEM of the functional connectivity change averaged across each cluster. Bar charts are included for explanatory purposes to show the nature of the interaction; magnitude should not be interpreted.

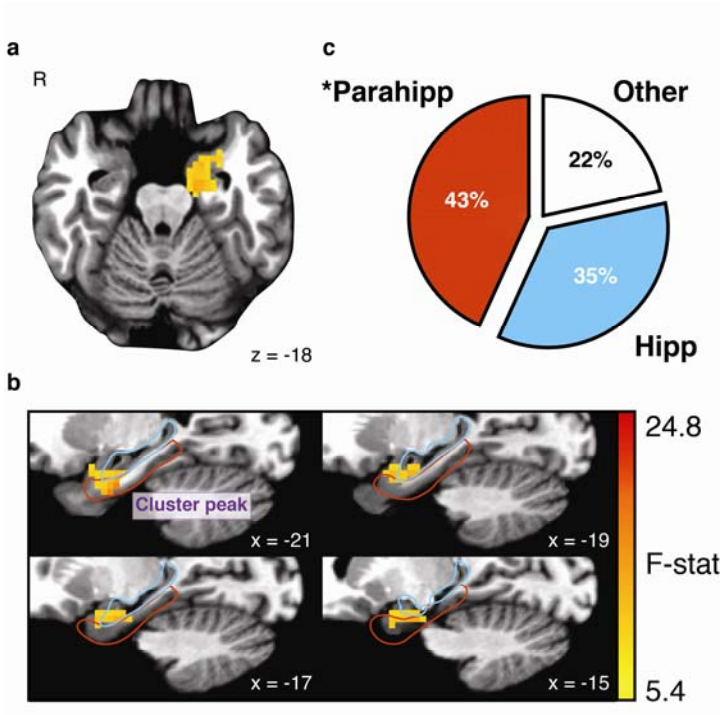

Figure 3. Impact of feedback during training on connectivity between premotor cortex and anterior MTL. In order to better examine the anatomical specificity of the region showing and interaction connectivity change due to training and feedback Group (Rest x Group interaction) of the MTL cluster, we calculated the proportion of significant voxels that fell within the hippocampus (blue) and parahippocampal cortex (red), defined by the Eickhoff-Zilles atlas. The peak voxel fell in parahippocampal cortex. Overall, approximately 78% of the voxels fell within the hippocampus and parahippocampal cortex. Voxels that did not fall within these regions were confined to the uncus and amygdala (white). This cluster reflects those voxels surviving correction for multiple comparisons (α = 0.05, $p < 0.005$, $k = 54$).

Feedback affects the relationship between PMC-cerebellar functional connectivity and 24-48h retention.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

In both tasks, all feedback groups showed robust sequence knowledge after learning; however, we hypothesized that the neural structures subserving this knowledge would be modulated by the type of feedback (Steel A, EH Silson, *et al.* 2016). To assess whether the changes in PMC connectivity were reflected in skill retention measured after 24 hours, we calculated the correlation between the change in functional connectivity due to training between PMC and the seven regions discovered in the previous analysis with skill retention at 24-48h.

Across the whole brain, the only region in which there was a feedback-dependent correlation between training-induced change in functional connectivity with the PMC and 24-48h skill retention was the cerebellum (Figure 4 upper): increased PMC-cerebellar connectivity change was positively associated with 24-48h skill retention in the REW ($r_{(24)}$=0.51, p<0.01) and PUN groups ($r_{(22)}$<0.54, p=0.012), but was negatively related to skill retention in the CONT group ($r_{(24)}$= -0.51, p<0.015). These correlation values were significantly different after correction for multiple comparisons [Figure 4a; corrected p=0.05/18=0.0027; REW v CONT (z=3.65, p<0.0001), PUN v CONT (z=3.68, p<0.0001), REW v PUN (z=0.13, p=0.89)]. PMC-cerebellar connectivity was not related to 1-h skill memory in any group, which suggests that this effect reflects overnight skill retention rather than motor performance (correlation values between functional connectivity change and 1-h skill retention: CONT: $r_{(24)}$<-0.10, p=0.65; REW: $r_{(24)}$<0.17, p=0.438; PUN: $r_{(24)}$<-0.006, p=0.978), although the difference between the strength of the correlation of functional connectivity change and 1-h and 24-48h did not reach statistical significance (CONT: z=1.98, p=0.13; REW: z=-1.267, p=0.21; PUN: z=1.672, p=0.058).

The relationship between PMC-caudate functional connectivity and 24-48h skill retention differs by feedback group.

Given the importance of the caudate in early sequence learning and feedback processing (Wachter T *et al.* 2009; Peterson EJ and CA Seger 2013), we expected to see differential relationship between PMC-caudate connectivity and 24-48h skill

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

memory across the feedback groups (Figure 4, lower). All groups showed an increase in connectivity between left PMC and left caudate after learning (LME, Main effect of Rest, $F_{(2,210)}$=52.119, p<0.001), and there was no difference in the magnitude of increase across the groups (LME, Rest x Group, $F_{(2,207)}$=1.575, p=0.21).

Despite the similarity of the connectivity change after training across the groups, the relationship between the connectivity change and 24-48h skill memory differed. Both the CONT and REW groups showed a strong correlation between left caudate and PMC connectivity change with skill memory at 24-48h after learning, such that greater connectivity was associated with better memory (CONT: $r_{(24)}$=0.73, p<0.0001; REW: $r_{(24)}$=0.60, p<0.003). In contrast, there was no relationship between the increase in PMC-caudate connectivity and 24-48h skill memory for the PUN group ($r_{(22)}$=0.07, p =0.46), which was significantly different from the CONT group (PUN v CONT: z=2.712, corrected p<0.0067) and the REW group (PUN v REW: z=-1.968, corrected p=0.049). PMC-caudate connectivity was not related to 1-h performance in any group, suggesting that this effect is specific to memory storage (CONT: $r_{(24)}$<-0.03, p=0.87; REW: $r_{(24)}$<0.39, p=0.071; PUN: $r_{(24)}$<-0.001, p=0.99); although only CONT showed a significant difference between the correlation at 1-h and 24-48h (CONT: z=3.107, p<0.005; REW, z=-0.912, p=0.36; PUN: z=0.165, p=0.87).

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

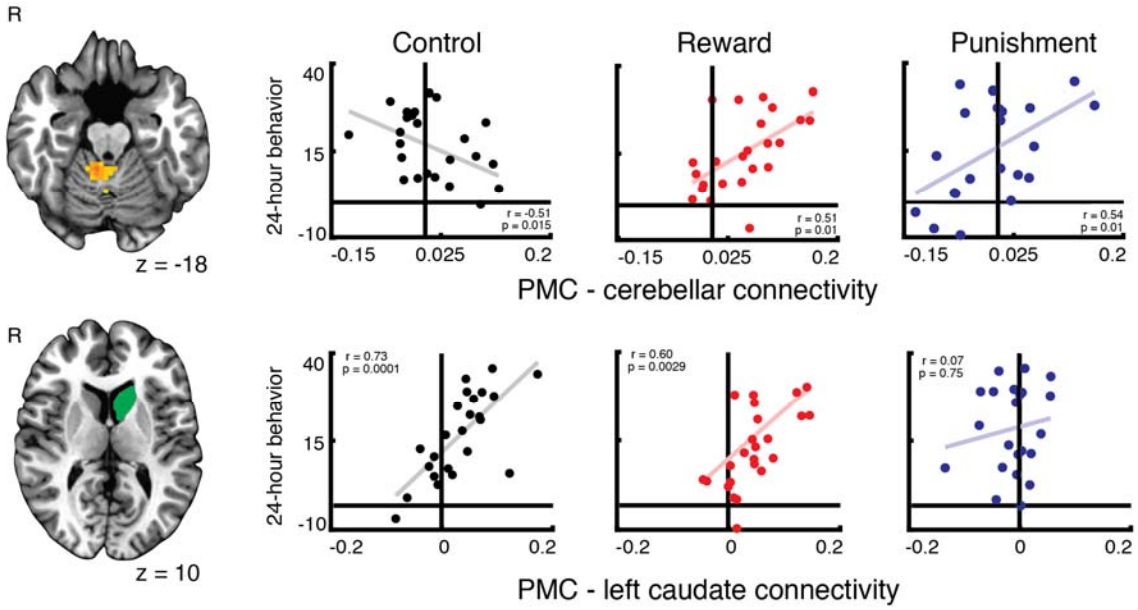

Figure 4. Feedback affects neural correlates of 24-48h skill retention. (Top) Increased PMC-medial cerebellum as defined by the interaction between functional connectivity after training and Feedback group was positively related to performance after training with REW or PUN, but was negatively related to performance after training with CONT feedback. (Bottom) Increased functional connectivity between left caudate (anatomically defined; green) and left PMC after training was positively related to 24-48h skill retention after training with REW and CONT feedback. Training with PUN broke this brain-behavior relationship. Scatter plots show correlation between 24-48h skill retention with average functional connectivity change from pre- to post- training when controlling for nuisance variables (motion and global correlation) for each feedback group.

Influence of feedback on correlation between PMC connectivity and 24-48h retention is specific to caudate and cerebellum.

So far, we have investigated whether connectivity of PMC predicted behavior in a feedback-dependent manner. Next, we adopted a whole-brain approach and tested whether the different effects of feedback group on the correlation between PMC connectivity change and 24-48h skill retention were specific to changes in PMC-cerebellar/caudate connectivity or were distributed across the brain. To this end, we added 24-48h skill retention as a covariate in the linear mixed effects model and

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

examined the interaction between Rest, Group, and 24-48h retention at each voxel. Memory at 1-h and 3+ weeks were included as covariates the model to control for task-performance effects and non-specific memory effects, respectively. Thus, the three-way interaction between Rest, Group, and 24-48h skill retention describes the behavioral relationship specific to the 24-48h probe. Confirming our previous finding using an ROI over the caudate, both the REW and CONT groups showed a strong positive relationship between increased PMC-left caudate connectivity and 24-48h memory, while no such relationship was demonstrated in the PUN group.

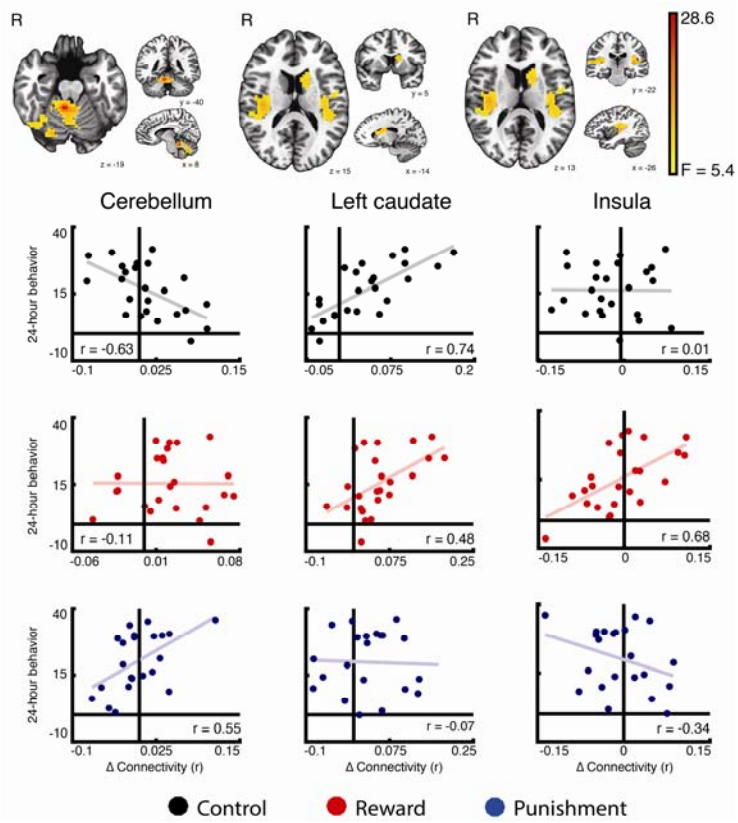

Figure 5. Whole–brain search for regions where correlation between functional connectivity change after training differed by feedback group. The relationship between 24-48h skill retention differed by Feedback group in cerebellum, left caudate, and insula. Scatter plots show correlation between 24-48h skill retention with average functional connectivity change from pre- to post- training when controlling for nuisance variables (motion and global correlation) for each feedback group. Notably, the cerebellar cluster that was significant at the whole brain level showed a different pattern of results than those

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

described in Figure 4. This is due to the larger spatial extent of the cluster found to be significant at the whole brain level, which had a heterogeneous relationship between 24-48h memory and connectivity change. This heterogeneity is further described in Figure 6. Scatter plots are included to show the nature of the interaction in each region, but the magnitude of the correlation should not be interpreted.

In addition to the correlations between PMC-caudate connectivity and behaviour, the relationship between left PMC and cerebellum connectivity to 24-48h skill memory retention was also affected by feedback. Notably, the regions of cerebellum which showed this interaction with behavior encompassed not only the dorsal medial portion of the cerebellum reported in the prior analyses above (see Figure 4), but also the ventromedial and dorsolateral cerebellum. When we examined the nature of this interaction across the whole of the cerebellar cluster, after training with PUN, connectivity between PMC and cerebellum was positively associated with skill at 24-48h (Figure 5), REW showed no relationship between connectivity and behavior, and CONT showed a negative relationship between connectivity and skill retention at 24-48h.

We hypothesized that the apparent disparity in these results was due to heterogeneity of the relationship between connectivity change and behavior across the cerebellum. We confirmed that this was the case; when we examined the individual local maxima within the cerebellar cluster (located in the dorsal medial, ventral medial, and dorsal lateral cerebellum) separately, we determined that the three spatial regions showed different relationships between connectivity change after training and 24-48h skill retention(Figure 5). The dorsal medial component, which overlaps with the region identified in the prior analyses reported above, showed a positive relationship with both REW and PUN (Figure 6, left). However, the CONT group showed a negative relationship between connectivity and 24-48h skill retention. This recapitulates our results from the previous ROI analysis. In contrast, the ventral medial cerebellum (Figure 6, middle) showed a weak positive correlation between functional connectivity change and 24-48h retention in the PUN group, but no correlation in the REW group and a negative correlation in the

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

CONT group. Finally, in the dorsolateral cerebellum (Figure 6, right), we found that the REW group had a negative relationship between functional connectivity and 24-48h skill retention, while the PUN group showed a weak positive correlation and the CONT group a negative correlation. [1]

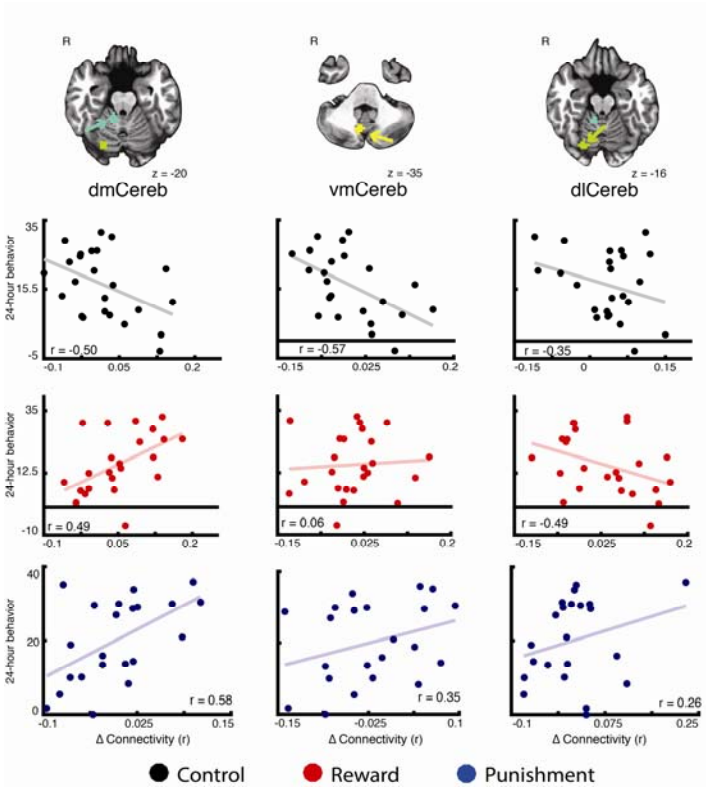

Figure 6. Relationship between premotor cortex-cerebellar connectivity change after training and skill retention at 24-48h differs across the cerebellum. In order to understand how connectivity between premotor cortex with different regions within the cerebellum were related to 24-48h skill retention, we separately analyzed each local maxima within the cerebellum where correlation between connectivity change with premotor cortex and 24-48h skill retention was determined to vary by group. The pattern of connectivity-behavior relationship differed across the cerebellum. In the dorsal medial cerebellum

---

[1] It is important to note that the aim of the analysis of the cerebellar subcomponents was to better understand the result seen at the whole-cluster level. Therefore, as the selection of these regions was based on a positive test on the same data, the magnitude of the relationships should not be interpreted.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

(left), which overlapped with the region of interest result show in Figure 4, both reward and punishment showed a strong positive relationship, and control showed a negative relationship; this recapitulates the result described in the ROI analysis (Figure 4). At the second local maxima, in the ventral medial cerebellum (middle), punishment showed a weak positive relationship with 24-48h retention, while reward showed no relationship. Again, control showed a negative relationship between connectivity and behavior. However, in the dorsal lateral cerebellum, punishment showed a weak positive relationship while reward showed a strong negative relationship. Control showed a negative relationship with all cerebellar regions. This spatial heterogeneity explains the difference in the result for the cerebellum at the whole brain level (Figure 5) compared to the region of interest (Figure 4). Scatter plots show the relationship between 24-48h skill retention and change in connectivity controlling for nuisance variables (motion and global correlation).

Finally, greater training-related connectivity between PMC and bilateral posterior insula was predictive of 24-48h memory in the REW group, while these connections were negatively predictive of 24-48h memory in the PUN group, and showed no predictive power in the CONT group.

Importantly, no regions showed differential correlations across the feedback groups between training-related changes in functional connectivity and 1-h or 3+ weeks retention test performance.

Discussion

In this study, we examined the effect of reward and punishment on neural mechanisms subserving skill memory. PMC connectivity was impacted by the type of feedback given during training: i) REW caused PMC connectivity with SMA and right dorsal medial cerebellum to increase; ii) PUN caused PMC connectivity with right anterior insula and left MTL to increase. Further, these connectivity changes due to training were behaviorally relevant. Specifically, in the PUN and REW groups, functional connectivity change between left PMC and right dorsal medial cerebellum correlated with skill memory at 24-48h, while for the CONT group connectivity

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

between left PMC and right dorsal medial cerebellum negatively correlated with skill memory. Finally, the relationship between PMC-caudate connectivity and behavior varied across the feedback groups. In the REW and CONT groups, increased PMC-caudate connectivity after training predicted greater 24-48h skill memory, but the PUN group showed no relationship. Collectively, reward and punishment may recruit separable neural resources, which could be exploited in rehabilitation or training regimes.

Our results may reflect the influence of valenced feedback on the learning systems recruited as a product of the learning strategy employed. Specifically, motivation by punishment encourages recruitment of hippocampus and lateral cerebellum, which are associated with a model-based learning strategy. In contrast, rewarded encourages recruitment of the caudate, which is associated with a model-free learning strategy. There is a long-standing proposal in the motor learning literature that both error-based and reinforcement feedback engage different neural mechanisms (Haith AM and JW Krakauer 2013). Several studies have demonstrated a behavioral dissociation between these two types of learning during visuomotor adaptation (Izawa J and R Shadmehr 2011; Shmuelof L *et al.* 2012; Galea JM *et al.* 2015; Cashaback JGA et al. 2017). In short, both model-free (or reinforcement-based) and model-based (or error-based) feedback are sufficient to learn a given task, but their learning properties make them more suitable for particular circumstances. Error-based learning integrates all possible learning exemplars (or trials), which affords a more generalizable model and allows learning to occur at a faster rate. However, this type of learning is computationally intensive, which may require sufficient motivation or even preclude the use of this approach in certain situations. In contrast, a learner using a model-free approach, based on reward prediction error, accumulates evidence on a trial-by-trial basis. Though slow, this approach is computationally easy and always results in optimal behavior (Haith AM and JW Krakauer 2013). Using a combination of transcranial direct current stimulation and transcranial magnetic stimulation, a dissociation was reported between the contribution of primary motor cortex-cerebellar connectivity and local

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

changes in primary motor cortex synaptic plasticity in the context of error-based and reinforcement-based learning, respectively (Uehara S et al. 2017).

Valenced feedback impacts three distinct PMC functional networks.

We found that PMC functional connectivity varied after training with reward, punishment, and control feedback. All groups showed equal 24-48h skill retention (Steel A, EH Silson, *et al.* 2016). Therefore, each of these networks may be sufficient to learn sequences, but feedback may bias the recruitment from one network to another. Previous studies reported functional connectivity increases within the motor system (Bassett DS et al. 2015) as well as between the parietal cortex, basal ganglia or MTL during skill automatization (Robertson EM et al. 2004; Debas K et al. 2014; Sami S et al. 2014). This disparity has been attributed to separable accounts of memory formation, such as model-based and model-free (Haith AM and JW Krakauer 2013), or implicit versus explicit memory (Sami S *et al.* 2014). Our result adds to this body of literature, suggesting that motivation can also impact the memory system engaged during the post-training period.

Effect of informative feedback on cerebellum.

We found that PMC-cerebellar connectivity differentially predicted 24-48h memory across feedback groups. CONT (uninformative feedback) showed a negative relationship between PMC-cerebellar functional connectivity change with 24-48h memory, while PUN showed a positive relationship. All clusters identified in the cerebellum were negatively related to behavior in CONT. The irrelevant exogenous feedback given to CONT may have impaired cerebellar processing, thereby degrading the quality of the model formed, and altering this relationship.

In REW the relationship of PMC-cerebellar connectivity differed across the cerebellum: PMC connectivity change with the dorsal lateral cerebellum was negatively related to 24-48h skill retention, while the ventral-medial and dorsal-medial aspects was positively related. In addition to its role in movement (Galea JM et al. 2011; Caligiore D et al. 2016) the cerebellum also contributes to cognitive

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

tasks (Krienen FM and RL Buckner 2009; Buckner RL 2013): the dorsolateral region differentially engaged by REW and PUN has greater functional connectivity with frontal cortex than the motor cortex (Krienen FM and RL Buckner 2009), reinforcing the hypothesis that REW engages motor regions while PUN engages executive regions.

Role of caudate in offline memory processing.

Functional connectivity between PMC and caudate increased after learning in all groups. However, PMC-caudate connectivity predicted 24-48h skill retention after training in CONT and REW, but not PUN. Others have found that head of caudate plays a key role in feedback processing during skill learning (Seger CA and CM Cincotta 2005; Peterson EJ and CA Seger 2013). For example, when learning non-motor skills, positive feedback elicits greater responses in caudate than negative feedback (Seger CA and CM Cincotta 2005). Caudate connectivity prior to motor learning also predicts REW-related memory (Hamann JM et al. 2014), and training on the SRTT with reward increases caudate activation (Wachter T *et al.* 2009). Our findings show that the contribution of the caudate to memory extends into the offline period, and that having trained with punishment may alter caudate processing after training.

Medial temporal lobe promotes rapid memory formation after punishment.

In our study, PMC-MTL connectivity increased after training with punishment. MTL is critical for both storage and recall of long-term memories. MTL contributes to in rapid, early sequence earning (Schendan HE *et al.* 2003; Albouy G et al. 2013) and representation (Ergorul C and H Eichenbaum 2006). We found that PMC-hippocampus connectivity increased after training in PUN, but not in REW or CONT. Hippocampus-dependent consolidation may require sleep [e.g. (Albouy G *et al.* 2013)], therefore, the observed connectivity increase immediately after training in PUN may not be directly related to consolidation, but rather preparatory activity in advance of sleep-dependent consolidation.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Connectivity between PMC and MTL may compensate for disrupted caudate processing in PUN. Alternatively, MTL and striatum, including caudate nucleus, may act competitively during learning (Poldrack RA et al. 2001). While we cannot distinguish between these possibilities, further work should investigate how the recruitment of the MTL after training with PUN impacts the quality of the memory formed.

Notably, we did not find a relationship between MTL functional connectivity change with 24-48h skill memory. It may be that the memory benefit from this enhanced connectivity requires sleep, or that this connectivity increase reflects enhanced encoding, rather than a memory retention related effect. However, the specificity of the decrease is compelling and warrants investigation in future studies.

Methodological considerations

Several aspects of our study are worth highlighting. First, we did not distinguish between dorsal and ventral premotor cortex in our ROI analysis because there was no Condition x ROI interaction. Premotor dorsal and ventral are highly interconnected. Although feedback may differentially impact dorsal and ventral premotor cortex, we may have been underpowered to detect this effect. Second, several regions where feedback impacted functional connectivity after training did not relate to behavior. While this complicates interpretation of our results, our study offers hypotheses that may be tested using interventional approaches. Third, our implementation of the SRT might foster explicit knowledge, which might recruit different neural networks after learning. This was necessary in order to accommodate fMRI acquisition during task performance. No participants included in the study spontaneously reported sequence knowledge. In order to prevent contamination of the participants organic experience, we did not make participants aware of the sequence until the final test session 3+ weeks after training. However, when tested at 3+ weeks, participants showed no evidence of explicit awareness (for further discussion, see Steel et al., 2016a).

Conclusions.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

We report that dissociable networks are recruited during the period immediately after training with REW, PUN, and CONT feedback: REW showed additional recruitment of the motor network, PUN recruited the hippocampal network, and CONT recruited an occipitotemporal network. In addition, the relationship between PMC-cerebellum, and PMC-caudate connectivity to 24-48h skill retention varied based on the type of feedback given during training. These results demonstrate the impact of feedback valence and elucidate potential mechanisms that may be exploited to enhance skill learning in the clinic.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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, 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.

Bassett DS, Yang M, Wymbs NF, Grafton ST. 2015. Learning-induced autonomy of sensorimotor systems. Nature neuroscience 18:744-751.

Buckner RL. 2013. The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. Neuron 80:807-815.

Caligiore D, Pezzulo G, Baldassarre G, Bostan AC, Strick PL, Doya K, Helmich RC, Dirkx M, Houk J, Jorntell H, Lago-Rodriguez A, Galea JM, Miall RC, Popa T, Kishore A, Verschure PF, Zucca R, Herreros I. 2016. Consensus Paper: Towards a Systems-Level View of Cerebellar Function: the Interplay Between Cerebellum, Basal Ganglia, and Cortex. Cerebellum.

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.

Cashaback JGA, McGregor HR, Mohatarem A, Gribble PL. 2017. Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. PLoS Comput Biol 13:e1005623.

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.

Eickhoff SB, Heim S, Zilles K, Amunts K. 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. NeuroImage 32:570-582.

Ergorul C, Eichenbaum H. 2006. Essential role of the hippocampal formation in rapid learning of higher-order sequential associations. The Journal of neuroscience : the official journal of the Society for Neuroscience 26:4111-4117.

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.

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.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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.

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

Galea JM, Vazquez A, Pasricha N, de Xivry JJ, Celnik P. 2011. Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns. Cerebral cortex 21:1761-1770.

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.

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.

Haith AM, Krakauer JW. 2013. Model-based and model-free mechanisms of human motor learning. Adv Exp Med Biol 782:1-21.

Hamann JM, Dayan E, Hummel FC, Cohen LG. 2014. Baseline frontostriatal-limbic connectivity predicts reward-based memory formation. Human brain mapping 35:5921-5931.

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.

Izawa J, Shadmehr R. 2011. Learning from sensory and reward prediction errors during motor adaptation. PLoS Comput Biol 7:e1002012.

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.

Krienen FM, Buckner RL. 2009. Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. Cerebral cortex 19:2485-2497.

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.

Murty VP, Labar KS, Adcock RA. 2012. Threat of punishment motivates memory encoding via amygdala, not midbrain, interactions with the medial temporal lobe.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

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.

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.

Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, Gluck MA. 2001. Interactive memory systems in the human brain. Nature 414:546-550.

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.

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.

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.

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

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Seger CA, Cincotta CM. 2005. The roles of the caudate nucleus in human classification learning. The Journal of neuroscience : the official journal of the Society for Neuroscience 25:2941-2951.

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.

Song S, Gotts SJ, Dayan E, Cohen LG. 2015. Practice structure improves unconscious transitional memories by increasing synchrony in a premotor network. Journal of cognitive neuroscience 27:1503-1512.

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.

Steel A, Song S, Bageac D, Knutson KM, Keisler A, Saad ZS, Gotts SJ, Wassermann EM, Wilkinson L. 2016. Shifts in connectivity during procedural learning after motor cortex stimulation: A combined transcranial magnetic stimulation/functional magnetic resonance imaging study. Cortex; a journal devoted to the study of the nervous system and behavior 74:134-148.

Stillman CM, Gordon EM, Simon JR, Vaidya CJ, Howard DV, Howard JH, Jr. 2013. Caudate resting connectivity predicts implicit probabilistic sequence learning. Brain connectivity 3:601-610.

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.

Uehara S, Mawase F, Celnik P. 2017. Learning Similar Actions by Reinforcement or Sensory-Prediction Errors Rely on Distinct Physiological Mechanisms. Cerebral cortex:1-13.

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.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Tables

Table1. Significant effects detected for premotor cortex linear mixed effects model. Clusters significant at p > 0.005, k=54

Premotor connectivity

Cluster Size (vxl)	Cluster Size (mm^3)	Peak x	Peak y	Peak z	Region (peak)
Post-training > Pre-training					
3719	100413	-4.5	-22.5	11.5	Thalamus
Pre-training > Post-training					
896	24192	52.5	1.5	-6.5	Right superior temporal gyus
154	4158	-58.5	-4.5	-0.5	Left superior temporal gyrus
82	2214	10.5	4.5	18.5	Right parahippocampal gyrus
66	1782	-4.5	-28.5	68.5	Left medial frontal gyrus
Rest x Group interaction					
204	5508	25.5	-73.5	38.5	Right IPL
184	4968	-19.5	-76.5	38.5	Left IPL
145	3915	7.5	-19.5	65.5	SMA
89	2403	-31.5	-70.5	-6.5	Left lateral occipitotemporal cortex
85	2295	43.5	4.5	-6.5	Right anterior insula
78	2106	7.5	-40.5	-21.5	Cerebellum
73	1971	-22.5	-10.5	-24.5	Left MTL
Rest x Group x 24-48h					
482	13014	7.5	-40.5	-18.5	Medial cerebellum
227	6129	-34.5	-22.5	5.5	Left posterior insula
216	5832	37.5	-19.5	14.5	Right posterior insula
93	2511	-13.5	-1.5	23.5	Left caudate

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Captions

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. Participants were then tested for skill memory 1-h, 24-48h, and 3+ weeks after initial training. 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. Premotor cortices connectivity change varies by feedback group. (*Right*) Linear mixed effects model revealed brain regions exhibiting a Rest x Feedback group interaction. (Left) These included parietoccipital regions (i), where connectivity increased after training with CONT feedback, supplementary motor area and cerebellum (ii), where connectivity increased after training with REW, and MTL and anterior insula (iii), where connectivity increased after training with PUN. Bar chart shows mean ± SEM of the functional connectivity change averaged across each cluster. Bar charts are included to show the nature of the interaction, but magnitude should not be interpreted.

Figure 3. Impact of feedback during training on connectivity between premotor cortex and anterior MTL. In order to better examine the anatomical specificity of the region showing and interaction connectivity change due to training and feedback Group (Rest x Group interaction) of the MTL cluster, we calculated the proportion of significant voxels that fell within the hippocampus (blue) and parahippocampal cortex (red), defined by the Eickhoff-Zilles atlas. The peak voxel fell in parahippocampal cortex. Overall, approximately 78% of the voxels fell within the hippocampus and parahippocampal cortex. Voxels that did not fall within these

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

regions were confined to the uncus and amygdala (white). This cluster reflects those voxels surviving correction for multiple comparisons ($\alpha = 0.05$, $p < 0.005$, $k = 54$).

Figure 4. Feedback affects neural correlates of 24-48h skill retention. (Top) Increased PMC-medial cerebellum as defined by the interaction between functional connectivity after training and Feedback group was positively related to performance after training with REW or PUN, but was negatively related to performance after training with CONT feedback. (Bottom) Increased functional connectivity between left caudate (anatomically defined; green) and left PMC after training was positively related to 24-48h skill retention after training with REW and CONT feedback. Training with PUN broke this brain-behavior relationship. Scatter plots show correlation between 24-48h skill retention with average functional connectivity change from pre- to post- training when controlling for nuisance variables (motion and global correlation) for each feedback group.

Figure 5. Whole–brain search for regions where correlation between functional connectivity change after training differed by feedback group. The relationship between 24-48h skill retention differed by Feedback group in cerebellum, left caudate, and insula. Scatter plots show correlation between 24-48h skill retention with average functional connectivity change from pre- to post- training when controlling for nuisance variables (motion and global correlation) for each feedback group. Notably, the cerebellar cluster that was significant at the whole brain level showed a different pattern of results than those described in Figure 4. This is due to the larger spatial extent of the cluster found to be significant at the whole brain level, which had a heterogeneous relationship between 24-48h memory and connectivity change. This heterogeneity is further described in Figure 6. Scatter plots are included to show the nature of the interaction in each region, but the magnitude of the correlation should not be interpreted.

bioRxiv preprint doi: https://doi.org/10.1101/243899; this version posted January 5, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Figure 6. Relationship between premotor cortex-cerebellar connectivity change after training and skill retention at 24-48h differs across the cerebellum. In order to understand how connectivity between premotor cortex with different regions within the cerebellum were related to 24-48h skill retention, we separately analyzed each local maxima within the cerebellum where correlation between connectivity change with premotor cortex and 24-48h skill retention was determined to vary by group. The pattern of connectivity-behavior relationship differed across the cerebellum. In the dorsal medial cerebellum (left), which overlapped with the region of interest result show in Figure 4, both reward and punishment showed a strong positive relationship, and control showed a negative relationship; this recapitulates the result described in the ROI analysis (Figure 4). At the second local maxima, in the ventral medial cerebellum (middle), punishment showed a weak positive relationship with 24-48h retention, while reward showed no relationship. Again, control showed a negative relationship between connectivity and behavior. However, in the dorsal lateral cerebellum, punishment showed a weak positive relationship while reward showed a strong negative relationship. Control showed a negative relationship with all cerebellar regions. This spatial heterogeneity explains the difference in the result for the cerebellum at the whole brain level (Figure 5) compared to the region of interest (Figure 4). Scatter plots show the relationship between 24-48h skill retention and change in connectivity controlling for nuisance variables (motion and global correlation).