

Covert skill learning in a cortical–basal ganglia circuit

Jonathan D. Charlesworth¹, Timothy L. Warren¹ & Michael S. Brainard¹

We learn complex skills such as speech and dance through a gradual process of trial and error. Cortical–basal ganglia circuits have an important yet unresolved function in this trial-and-error skill learning¹; influential ‘actor–critic’ models propose that basal ganglia circuits generate a variety of behaviours during training and learn to implement the successful behaviours in their repertoire^{2,3}. Here we show that the anterior forebrain pathway (AFP), a cortical–basal ganglia circuit⁴, contributes to skill learning even when it does not contribute to such ‘exploratory’ variation in behavioural performance during training. Blocking the output of the AFP while training Bengalese finches to modify their songs prevented the gradual improvement that normally occurs in this complex skill during training. However, unblocking the output of the AFP after training caused an immediate transition from naive performance to excellent performance, indicating that the AFP covertly gained the ability to implement learned skill performance without contributing to skill practice. In contrast, inactivating the output nucleus of the AFP during training completely prevented learning, indicating that learning requires activity within the AFP during training. Our results suggest a revised model of skill learning: basal ganglia circuits can monitor the consequences of behavioural variation produced by other brain regions and then direct those brain regions to implement more successful behaviours. The ability of the AFP to identify successful performances generated by other brain regions indicates that basal ganglia circuits receive a detailed efference copy of premotor activity in those regions. The capacity of the AFP to implement successful performances that were initially produced by other brain regions indicates precise functional connections between basal ganglia circuits and the motor regions that directly control performance.

We assessed the contributions of basal ganglia circuitry to learned modification of adult Bengalese finch song, a complex behaviour consisting of a sequence of 30–100-ms ‘syllables’, each with a highly stereotyped acoustic structure. The song-specific motor control system consists of a motor pathway, which is analogous to mammalian premotor and primary motor cortex and is sufficient to produce well-learned elements of song, and the AFP, which is necessary for juvenile song learning and adult song modification⁴. We elicited learning by training birds with aversive reinforcement contingent on the fundamental frequency of individually targeted syllables (Fig. 1a, b). Aversive reinforcement consisted of loud, 50–80-ms bursts of white noise^{5,6}. Training with aversive reinforcement caused songbirds to modify fundamental frequency in a direction that adaptively reduced the likelihood of white noise exposure; delivering white noise to performances of a syllable with fundamental frequency below a threshold caused an increase in mean fundamental frequency of that syllable (Fig. 1b), whereas delivery of white noise to performances with fundamental frequency above that threshold caused a decrease in mean fundamental frequency. These adaptive changes developed within hours and were specific to the fundamental frequency of the targeted syllable.

Influential actor–critic models^{2,3}, inspired by reinforcement learning theory⁷ and supported by empirical evidence^{8,9}, propose that basal ganglia circuits such as the AFP are a crucial substrate for trial-and-error learning, generating a variety of behavioural performances and

ultimately implementing only the performances that have led to successful outcomes. In the context of fundamental-frequency modification (Fig. 1a, b), the actor–critic model proposes that on each trial the AFP (the actor) generates distinct fundamental frequency values (exploratory behavioural variation; Fig. 1c), receives reinforcement

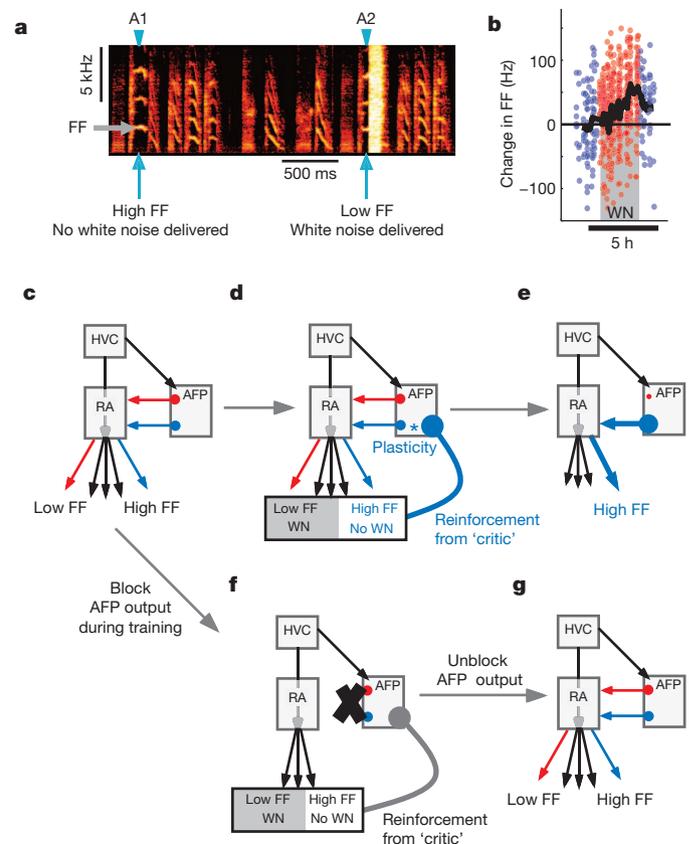


Figure 1 | Trial-and-error learning in adult birdsong. **a**, Spectrogram of song during an experiment in which white noise (WN) was delivered to targeted syllable (A) renditions with low fundamental frequency (FF) but not high FF. **b**, Delivering WN to syllables with low FF (shaded region) elicited increases in FF. Each point corresponds to one syllable rendition; the black line indicates the running average. **c**, The song circuit includes a motor pathway, containing HVC and RA, and the AFP, which is important for learning. The AFP generates variation in performance (motor exploration); red and light blue indicate distinct activity patterns in the AFP that lead to distinct FF values on different renditions of the same syllable. **d**, Actor–critic models propose that the AFP receives feedback about the behavioural variants that it generates, and this feedback strengthens patterns of AFP activity yielding better outcomes (light blue, feedback shown) and weakens patterns of AFP activity yielding worse outcomes (red). **e**, This changes the output of the AFP so that it selectively implements more successful behaviours. **f**, We tested this model by blocking the output of the AFP during training, thus preventing the AFP from generating variation in FF. **g**, The model predicts that this will prevent learning-related plasticity in the AFP, and thus there will be no change in FF, even when AFP output is unblocked after training.

¹W. M. Keck Center for Integrative Neuroscience, Department of Physiology, and the Neuroscience Graduate Program, University of California, San Francisco, California 94143, USA.

signals about the consequences of that variation from dopaminergic neurons (the critic; Fig. 1d), and changes the probability of generating that fundamental frequency value in the future on the basis of its consequences^{4,10–12}. Over time, the AFP gradually adjusts its output to implement (that is, to cause the execution of) behaviours with better consequences, leading to adaptive changes in fundamental frequency and thus improved skill performance (Fig. 1e). Consistent with this model, blocking AFP output through lesions or reversible inactivations reduces song variation, indicating that the AFP generates variation in song performance that might serve as motor exploration^{4,5} (Fig. 1c, f). Moreover, blocking AFP output after learning reduces the expression of recently learned song changes, suggesting that the AFP can contribute to learning by biasing the motor pathway to implement more successful behaviours^{13,14} (as suggested in Fig. 1e). A critical yet untested proposition of this model is that learning requires the reinforcement of exploratory behavioural variation generated by the AFP; therefore, preventing the AFP from contributing to behavioural variation during training should prevent trial-and-error learning (Fig. 1f, g).

We tested this prediction by pharmacologically blocking the output of the AFP, training birds with aversive reinforcement, and then unblocking the output of the AFP. To block contributions of the AFP to exploratory variation in song during training, while leaving intrinsic AFP circuitry intact, we exploited a pharmacological distinction between inputs that the songbird motor cortical nucleus RA (robust nucleus of the arcopallium) receives from premotor cortical nucleus HVC and from AFP output nucleus LMAN (lateral magnocellular nucleus of the anterior nidopallium). Inputs from LMAN are mediated almost exclusively by *N*-methyl-D-aspartate (NMDA) receptors whereas inputs from HVC are mediated by both NMDA receptors and α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptors⁴ (Fig. 2a). To disrupt AFP output reversibly we therefore inserted microdialysis probes into RA and used retrodialysis to switch between a control solution (artificial cerebrospinal fluid; ACSF) and a solution containing the NMDA receptor antagonist 2-amino-5-phosphonovaleric acid (APV) at 1–5 mM (Fig. 2a). Consistent with previous reports^{14,15}, this manipulation affected song in the same manner as pharmacological inactivations or lesions of LMAN^{14,16}, reducing the coefficient of variation (CV) of the fundamental frequency by $31.7 \pm 5.6\%$ ($n = 12$ syllables in 9 birds) without causing systematic changes in song structure (Fig. 2b, c and Supplementary Fig. 2). The APV-dependent reduction in song variation was reversible; switching the infusion solution back to ACSF restored the CV of the fundamental frequency to $96.5 \pm 4.6\%$ of baseline (Fig. 2c and

Supplementary Fig. 2c). These data indicate that infusing APV into RA effectively and reversibly prevents the AFP from contributing to song variation (as shown schematically in Fig. 1c, f).

As predicted by an actor–critic model of AFP function, there was no expression of learning while AFP output was blocked during training. We compared learning in control experiments (an example is shown in Fig. 3a) with learning in experiments with APV in RA throughout training (an example is shown in Fig. 3c). Training consisted of administering aversive reinforcement contingent on the fundamental frequency of a targeted syllable (Fig. 1a, b). To ensure that a similar proportion of syllable renditions received aversive reinforcement across experiments despite the reduced range of variation after APV infusion, we set the threshold for avoiding white noise at roughly the baseline median fundamental frequency for each targeted syllable (see Methods). To simplify presentation, we have plotted data so that the direction of learning (that reduces white noise exposure) is always upwards. For control experiments ($n = 14$ experiments for 9 syllables in 7 birds), there was significant expression of learning during the training period; the mean shift of fundamental frequency in the adaptive direction was 33.5 Hz, corresponding to a $1.1 \pm 0.35\%$ change in fundamental frequency (Fig. 3b, left bar; $P < 0.01$, signed-rank test). In contrast, for experiments with APV in RA ($n = 21$ experiments for 12 syllables in 9 birds), there was no expression of learning during the training period (Fig. 3d, left bar); the mean shift in fundamental frequency was 5.3 Hz (a $0.20 \pm 0.15\%$ change) which was significantly less than in control conditions ($P = 0.02$, rank-sum test) and not significantly different from zero ($P = 0.15$, signed-rank test). These results indicate that infusing APV into RA eliminates any expression of learning during training and thus provide further support that this manipulation blocks AFP output.

Learned changes to song appeared immediately when AFP output was unblocked after training. If learning required the AFP to transmit song variation during training, as predicted by an actor–critic model of AFP function, then blocking AFP output during training should have prevented learning and thus unblocking AFP output after training should not have revealed any learned changes to fundamental frequency (Fig. 1f, g). Contrary to this prediction, we observed learned changes to fundamental frequency after unblocking AFP output (Fig. 3c, d). These learned changes could not be predicted by any subtle changes in fundamental frequency during training (Supplementary Fig. 3) and were specific to the fundamental frequency of the targeted syllable (Fig. 3e and Supplementary Fig. 4). The average learned change across experiments was 27.6 Hz, corresponding to a

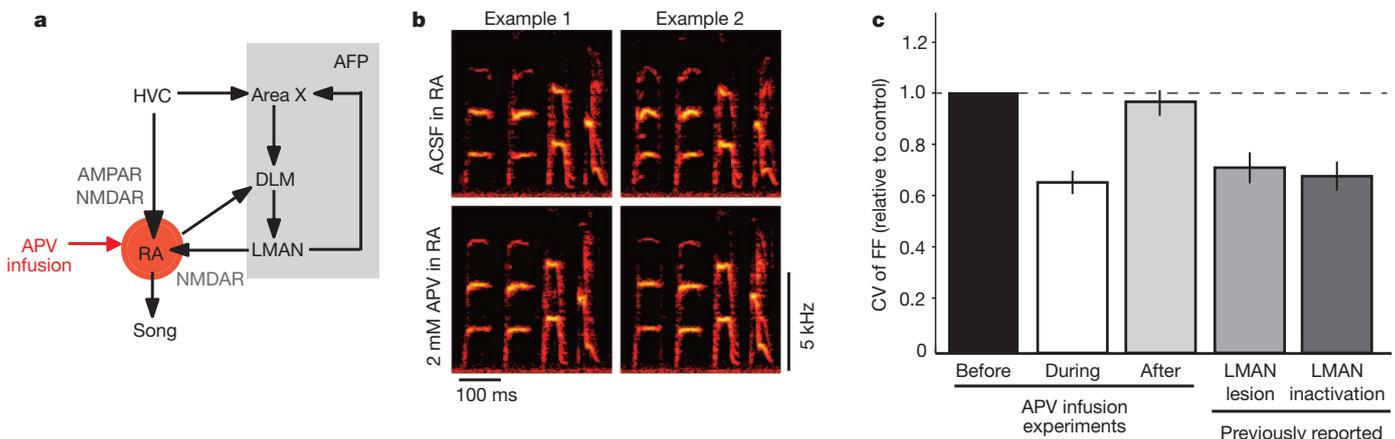


Figure 2 | Infusing APV into RA reduced song variability reversibly without distorting song structure. **a**, The AFP contains the striatopallidal nucleus Area X, the thalamic nucleus DLM and the cortical nucleus LMAN, which projects to RA. We blocked AFP output to the motor pathway by infusing the NMDA receptor (NMDAR) antagonist APV into RA. AMPAR, AMPA receptor. **b**, Infusion of APV into RA did not markedly change the song.

c, Infusions of APV into RA reduced the coefficient of variation (CV) of FF, which recovered after switching back to ACSF ($n = 12$ syllables in 9 birds). The decrease in CV with APV in RA ($31.7 \pm 5.6\%$) was not significantly different from previously reported effects of lesions ($34.1 \pm 4.5\%$) and inactivations ($28.4 \pm 6.0\%$) of LMAN in adult Bengalese finches. Error bars indicate s.e.m. Previously reported values are from refs 14 and 16.

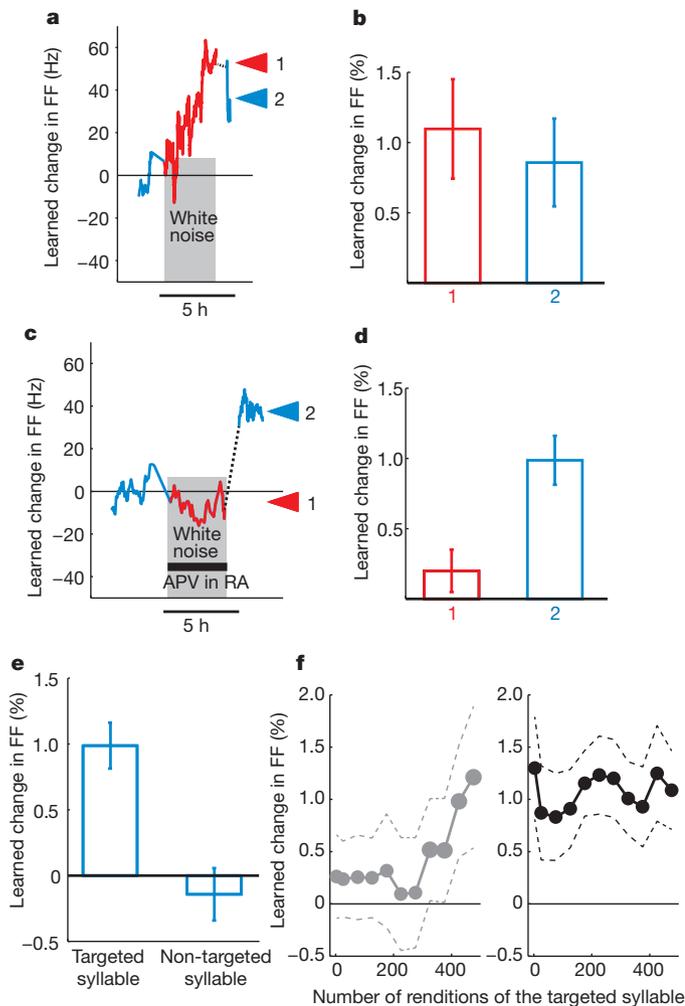


Figure 3 | Infusing APV into RA prevents the expression but not the acquisition of learning. **a, b,** Control experiments (ACSF in RA). **a,** Example of experiment in which white noise was delivered to targeted syllables with low FF. Arrowheads indicate FF at end of training (1) and after training (2). The dashed line indicates the delay between measurements at the end of training and after training. **b,** For control experiments ($n = 14$ experiments in 7 birds), learning was expressed at a similar magnitude at the end of training (1) and after training (2). Learning was normalized as a percentage of baseline FF. Error bars indicate s.e.m. **c, d,** Experiments with APV infused into RA. **c,** Example of experiment with AFP output blocked throughout the training period. Arrowheads indicate FF at end of training (1) and after training and APV washout (2). **d,** For experiments with APV in RA ($n = 21$ experiments in 9 birds), learning at end of training (1) was not significantly greater than zero and was significantly less than in control experiments. Learning after training and APV washout (2) was significantly greater than zero and was the same magnitude as in control experiments. **e,** After training and APV washout, learning was evident in syllables targeted with reinforcement (left) but not in other syllables of the same songs that were not targeted with reinforcement (right). This analysis was performed for each experiment in which FF of a non-targeted syllable could be reliably quantified ($n = 17$ of 21 total experiments). **f,** Mean progression of learning for control experiments (left) and after unblocking AFP output for experiments with APV in RA (right). Points correspond to syllable renditions 1–5, 1–50, 51–100, ..., 451–500. Dashed lines indicate s.e.m.

$0.99 \pm 0.17\%$ change in fundamental frequency ($n = 21$ experiments in 9 birds; Fig. 3d, right bar; $P < 0.001$, signed-rank test). The magnitude of learning expressed after training was statistically indistinguishable from the magnitude of learning in control experiments (Fig. 3b, d, right bars; $P > 0.9$, rank-sum test). In contrast to the gradual progression of learning in control experiments, maximal learning was expressed immediately after unblocking AFP output

and did not require further practice with AFP output unblocked (Fig. 3f). Thus, during training with AFP output blocked, the AFP had not only encoded a ‘policy’ specifying the change in song that would improve outcomes (for example, the fundamental frequency of the targeted syllable should be increased) but had already altered its activity to implement that change.

The acquisition of learning during training with APV in RA is consistent with three classes of mechanism. First, learning could require activity in the AFP during training. Second, learning could require plasticity upstream of the AFP, possibly in the ventral tegmental area, and the AFP could merely serve as a conduit between the site of plasticity and behavioural output. Third, learning could require plasticity downstream of the AFP, in RA, but the expression of that learning could be gated by AFP output¹⁴. To discriminate between these possible mechanisms we inactivated LMAN during training, by infusing muscimol ($n = 12$ experiments in 3 birds) or lidocaine ($n = 2$ experiments in 1 bird) into LMAN (Fig. 4a). Whereas infusing APV into RA blocked AFP output while leaving activity in the AFP intact, inactivating LMAN not only blocked AFP output but also disrupted activity within the AFP.

We found that activity in LMAN during training is crucial for learning. Inactivating LMAN reversibly reduced variation in fundamental frequency by the same amount as lesions of LMAN or infusion of APV into RA (CV decrease of $31.2 \pm 6.5\%$, $n = 14$; Supplementary Fig. 2b). We ensured in each case that the threshold for reinforcement continued to provide a directed instructive signal during the training period despite the reduced range of fundamental frequency variation (as in APV experiments; see Methods)⁶. As with infusing APV into RA, inactivating LMAN prevented any expression of learning during training; expression of learning during training with LMAN inactivated was $-0.19 \pm 0.37\%$ ($n = 14$, $P = 0.9$, signed-rank test) in comparison with $0.90 \pm 0.09\%$ ($n = 14$, $P = 1.2 \times 10^{-4}$, signed-rank test) in control experiments (Fig. 4b–d). However, in contrast to experiments with APV in RA, inactivation of LMAN during training prevented any acquisition of learning as assessed after the washout of drug ($-0.07 \pm 0.21\%$, $n = 14$, $P = 0.95$, signed-rank test; Fig. 4c, d). These results demonstrate that inactivating AFP nucleus LMAN during training prevents the acquisition of learning, and therefore that activity within the AFP during training is essential for learning.

Taken together, our results indicate that the capacity to adaptively modify a complex motor skill developed within the AFP during training with AFP output blocked. The prevention of learning by inactivating LMAN during training indicates that activity in the AFP is required for learning (Fig. 4). The immediate transition from naive performance to learned performance when we unblocked AFP output after training (Fig. 3) demonstrates that, during training, the AFP had gained the ability to improve behaviour even though that improvement was not yet expressed. For simpler forms of conditioning^{17,18}, such covert learning, indicating learning-related plasticity in the brain that is not accompanied by behavioural improvement, would only require that the brain region involved in learning received coarse signals about actions and stimuli¹⁹. In contrast, our results indicate that the brain region involved in learning, the AFP, receives detailed information (an efference copy²⁰) about the precise dynamics and timing of behavioural performance from the other brain regions controlling that performance.

Our results motivate a revision to models of song plasticity^{10–12} and influential actor–critic models of skill learning^{2,3}, which propose that essential learning-related signals develop only in brain regions that are ‘acting’ (that is, controlling behaviour). In contrast, our results indicate that the essential learning-related signals necessary to adaptively bias behaviour can develop in a basal ganglia circuit, the AFP, while it is prevented from contributing to behavioural performance and motor exploration. This indicates that motor exploration (that is, variation) generated by the AFP is not necessary for learning, and therefore a source of variation independent of the AFP can be exploited for

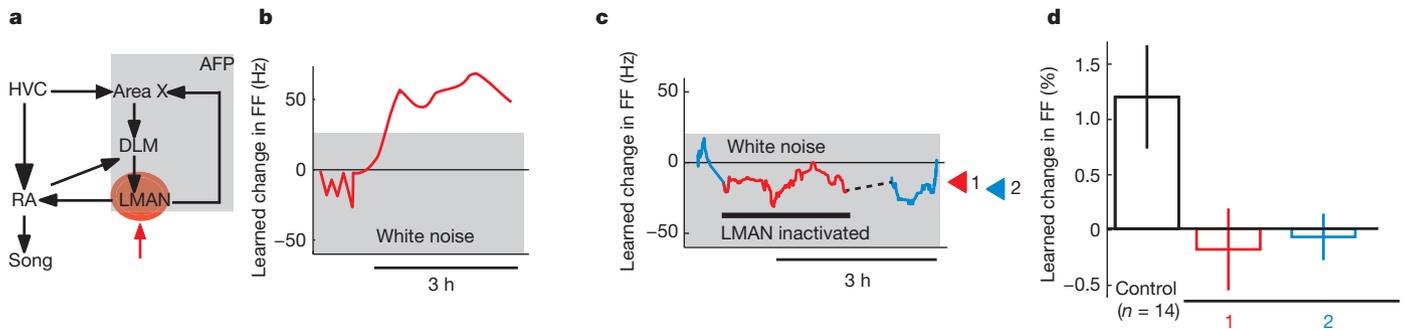


Figure 4 | Inactivating LMAN during training prevents both the expression and the acquisition of learning. **a**, We inactivated LMAN by infusing the GABA_A antagonist muscimol ($n = 12$ experiments in 3 birds) or the sodium channel blocker lidocaine ($n = 2$ experiments in 1 bird) into LMAN (red arrow). **b**, **c**, Example experiments. **b**, Control experiment in which white noise was delivered to renditions of a targeted syllable with low FF. **c**, As in **b**, but with

reinforcement learning. Presumably, this variation arises in the motor pathway, possibly in RA^{21,22}, and is transmitted to the AFP. In normal circumstances with AFP output intact, variation contributed by the AFP itself may also be used for reinforcement learning. The AFP may therefore be a specialized hub where information about behavioural variation from multiple sources converges and is associated with reinforcement signals to guide learning.

The specificity of learning with AFP output blocked (Fig. 3e and Supplementary Fig. 4) implies that the AFP associates reinforcement signals with detailed information about ongoing song performance, including both the identity of the syllable being produced and the rendition-by-rendition variation in the fundamental frequency of that syllable. Reinforcement signals, indicating the presence or absence of white noise, could be conveyed to the AFP by means of known projections from neuromodulatory nuclei such as the ventral tegmental area^{4,10}. Signals encoding syllable identity are conveyed to the AFP by means of projections from nucleus HVC in the motor pathway to the striatopallidal nucleus Area X (ref. 4). In principle, auditory feedback could provide information about variation in fundamental frequency, but such auditory signals seem to be absent from the AFP during singing²³. We therefore favour the alternative possibility, that information about fundamental frequency variation is transmitted to the AFP through an efference copy of activity in premotor regions, by way of projections from HVC to Area X and/or projections from RA to the basal ganglia-recipient thalamic nucleus DLM (dorsolateral division of the medial thalamus)^{24,25} (Supplementary Fig. 1). This is consistent with a recent proposal that the transmission of efference copy signals from motor cortex (HVC and/or RA) to basal ganglia circuitry (AFP) has a fundamental function in mammalian skill learning²⁶.

Our results also indicate precise functional coordination between the AFP and the motor pathway. Immediately after unblocking AFP output, we observed learning that was specific to the reinforced features of song, indicating that the AFP had modified its output to direct the production of those specific features by the motor pathway. This implies not only that the AFP receives detailed information about the song performances produced by the motor pathway during training, but also that it changes its output to specifically implement the features of those performances that were reinforced. Such a capacity of the AFP to precisely monitor and modify the activity of the motor pathway indicates fine-scale functional coordination both in the projections from the motor pathway to the AFP and in the projections from the AFP back to the motor pathway. Such bi-directional coordination might be mediated by segregated functional loops between the AFP and the motor pathway, each encoding a particular feature of song, such as high fundamental frequency in a particular syllable (Supplementary Fig. 1). Under normal conditions, with AFP output

intact, such functional loops could enable the AFP to amplify and bias specific behavioural features, functions that have been attributed to mammalian basal ganglia circuits^{27,28}. More generally, our results suggest that precise functional coordination between motor cortex and basal ganglia circuitry is important for enabling motor skill learning.

LMAN inactivated during training. Arrowheads indicate FF at the end of training with LMAN inactivated (1) and after training and muscimol washout (2). **d**, Summary: for experiments with LMAN inactivated (1 and 2; $n = 14$), there was neither evidence for learning at the end of training (red) nor after training and drug washout (light blue). Error bars indicate s.e.m.

METHODS SUMMARY

All experiments were performed on adult (more than 120 days old) male Bengalese finches (*Lonchura striata domestica*) singing undirected song. Song recording and feedback delivery were performed with software⁵ that recognized a targeted syllable and delivered a 50–80-ms burst of white noise unless the fundamental frequency (FF) met an escape criterion. For experiments with APV in RA and associated controls, the threshold for escaping white noise was set near the median FF of the targeted syllable; thus, about 50% of syllable performances initially avoided white noise. We used reverse microdialysis¹⁴ to deliver the NMDA-receptor antagonist DL-APV (1–5 mM in ACSF) to RA, and the GABA_A agonist muscimol (100–500 μ M) or the sodium channel blocker lidocaine (2%) to LMAN. To ensure the complete wash-in of drug, we delayed 1–2 h between drug infusion and the beginning of the training period. Immediately after training, the solution was switched back to ACSF. To ensure complete the wash-out of drug, we delayed at least 1 h between switching the solution to ACSF and measuring FF performance after training.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 October 2011; accepted 22 March 2012.

Published online 20 May 2012.

- Hikosaka, O., Nakamura, K., Sakai, K. & Nakahara, H. Central mechanisms of motor skill learning. *Curr. Opin. Neurobiol.* **12**, 217–222 (2002).
- Houk, J. C., Adams, J. L. & Barto, A. G. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.) 249–270 (MIT Press, 1995).
- Suri, R. E. & Schultz, W. A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* **91**, 871–890 (1999).
- Mooney, R. Neural mechanisms for learned birdsong. *Learn. Mem.* **16**, 655–669 (2009).
- Tumer, E. C. & Brainard, M. S. Performance variability enables adaptive plasticity of ‘crystallized’ adult birdsong. *Nature* **450**, 1240–1244 (2007).
- Charlesworth, J. D., Tumer, E. C., Warren, T. L. & Brainard, M. S. Learning the microstructure of successful behavior. *Nature Neurosci.* **14**, 373–380 (2011).
- Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
- Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
- Reynolds, J. N., Hyland, B. I. & Wickens, J. R. A cellular mechanism of reward-related learning. *Nature* **413**, 67–70 (2001).
- Fee, M. S. & Goldberg, J. H. A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience* **198**, 152–170 (2011).
- Fiete, I. R., Fee, M. S. & Seung, H. S. Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. *J. Neurophysiol.* **98**, 2038–2057 (2007).
- Doya, K. & Sejnowski, T. in *The New Cognitive Neurosciences* (ed. Gazzaniga, M.) 469–482 (MIT Press, 2000).

13. Andalman, A. S. & Fee, M. S. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proc. Natl Acad. Sci. USA* **106**, 12518–12523 (2009).
14. Warren, T. L., Tumer, E. C., Charlesworth, J. D. & Brainard, M. S. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *J. Neurophysiol.* **106**, 1806–1821 (2011).
15. Olveczky, B. P., Andalman, A. S. & Fee, M. S. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol.* **3**, e153 (2005).
16. Hampton, C. M., Sakata, J. T. & Brainard, M. S. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. *J. Neurophysiol.* **101**, 3235–3245 (2009).
17. Krupa, D. J., Thompson, J. K. & Thompson, R. F. Localization of a memory trace in the mammalian brain. *Science* **260**, 989–991 (1993).
18. Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W. & O'Reilly, R. C. Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nature Neurosci.* **10**, 126–131 (2007).
19. Balleine, B. W. & Ostlund, S. B. Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann. NY Acad. Sci.* **1104**, 147–171 (2007).
20. Crapse, T. B. & Sommer, M. A. Corollary discharge across the animal kingdom. *Nature Rev. Neurosci.* **9**, 587–600 (2008).
21. Olveczky, B. P., Otchy, T. M., Goldberg, J. H., Aronov, D. & Fee, M. S. Changes in the neural control of a complex motor sequence during learning. *J. Neurophysiol.* **106**, 386–397 (2011).
22. Sober, S. J., Wohlgenuth, M. J. & Brainard, M. S. Central contributions to acoustic variation in birdsong. *J. Neurosci.* **28**, 10370–10379 (2008).
23. Leonardo, A. Experimental test of the birdsong error-correction model. *Proc. Natl Acad. Sci. USA* **101**, 16935–16940 (2004).
24. Vates, G. E., Vicario, D. S. & Nottebohm, F. Reafferent thalamo-'cortical' loops in the song system of oscine songbirds. *J. Comp. Neurol.* **380**, 275–290 (1997).
25. Goldberg, J. H. & Fee, M. S. A cortical motor nucleus drives the basal ganglia-recipient thalamus in singing birds. *Nature Neurosci.* **15**, 620–627 (2012).
26. Redgrave, P. & Gurney, K. The short-latency dopamine signal: a role in discovering novel actions? *Nature Rev. Neurosci.* **7**, 967–975 (2006).
27. Turner, R. S. & Desmurget, M. Basal ganglia contributions to motor control: a vigorous tutor. *Curr. Opin. Neurobiol.* **20**, 704–716 (2010).
28. Frank, M. J. Computational models of motivated action selection in corticostriatal circuits. *Curr. Opin. Neurobiol.* **21**, 381–386 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Frank, A. Doupe, M. Stryker and D. Mets for discussion and comments on the manuscript. This work was supported by National Institutes of Health grant NIDCD R01 and National Institute of Mental Health grant P50. J.D.C. and T.L.W. were supported by National Science Foundation graduate fellowships.

Author Contributions J.D.C., T.L.W. and M.S.B. designed the experiments. J.D.C. performed the experiments with APV in RA, and T.L.W. performed the experiments with LMAN inactivations. J.D.C. analysed the data. J.D.C. prepared the manuscript, with input from the other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.D.C. (jcharles@phy.ucsf.edu).

METHODS

Animal care. All experiments were performed on adult (more than 120 days old) male Bengalese finches (*Lonchura striata domestica*) that had been bred in our colony and housed with their parents until at least 60 days of age. During experiments, birds were housed individually in sound-attenuating chambers (Acoustic Systems) with food and water provided *ad libitum*. All song recordings were from undirected song (that is, no female was present). All procedures were performed in accordance with established protocols approved by the University of California, San Francisco Institutional Animal Care and Use Committee.

Training. The same training parameters were used for control experiments and experiments with pharmacological manipulations. Song acquisition and feedback delivery were accomplished using previously described LabView software (EvTaf⁶), which recognized a specific time (contingency time) in a targeted syllable of song based on its spectral profile. On recognition, EvTaf recorded the time and calculated the fundamental frequency (FF) during the previous 8 ms of song. If the FF met the escape criterion (that is, above or below a threshold), no disruptive feedback was delivered. Otherwise, a 50–80-ms burst of white noise was delivered starting less than 1 ms after the contingency time. The duration of white noise was constant for a given experiment. To allow quantification of FF during training, a randomly interleaved 10% of songs were allocated as catch trials and did not receive white noise.

Experiments with reversible disruption of LMAN transmission to RA by reverse microdialysis. We interfered with LMAN transmission to RA by using a previously described reverse microdialysis technique¹⁴, in which solution diffuses into targeted brain areas across the dialysis membranes of implanted probes. RA was mapped electrophysiologically during cannula implantation so as to direct probes to the centre of RA. Between probe insertion and white noise training, there was a more than 48 h period in which control solution (ACSF) was dialysed at a flow rate of $1 \mu\text{l min}^{-1}$. The dialysis solution was switched from ACSF to the NMDA-receptor antagonist DL-APV (2–5 mM in ACSF; Ascent) at least 1.5 h before the onset of white noise training so that the threshold for escaping white noise could be determined on the basis of song performance with APV in RA. During this period we evaluated the efficacy of APV by assessing the rendition-to-rendition variability of FF for individual syllables. FF variability reduced and stabilized at an asymptotic level within the first 30 min of APV dialysis, indicating rapid onset and equilibrium of drug effect. We observed a reduction in variability similar to that reported after lesions or inactivations of LMAN^{14,16}. For clarity of presentation in Fig. 3, running averages of FF performance for experiments with APV in RA omit the period during APV wash-in before the onset of white noise. For experiments with APV in RA and the accompanying control experiments, white noise was delivered for 4–14 h while birds were awake. Blocking AFP output reduced variation in FF by an average of 31.7%, meaning that setting the threshold for avoiding white noise at a certain level above mean FF (for example +30 Hz) in control experiments and experiments with AFP output blocked would result in a greater proportion of syllable performances escaping aversive reinforcement in control experiments. To avoid this confound and ensure that a similar proportion of syllable renditions received aversive reinforcement in control experiments and experiments with AFP output blocked, we set the threshold for avoiding white noise at approximately the baseline median FF performance (between the 40th and 60th centiles in all experiments). To ensure that our assessment of learning during the training period evaluated the effects of white noise training as opposed to the acute effects of APV, FF change at the end of the training period was quantified by subtracting FF immediately before training (during the period with APV in RA before the onset of white noise) from FF at the end of the training period. Immediately after the conclusion of white noise training, the dialysis solution was switched back to ACSF. Learning after the training period was quantified by measuring the difference between FF performance after white noise training (with ACSF in RA) and FF performance before white noise training and before infusing APV into RA (that is, with ACSF in RA). Although the latency between switching the solution remotely at the pumping apparatus and changing the solution at the probe tips was only 6 min in our experimental setup¹⁴, the APV-dependent decrease in FF variability typically remained for hours after switching back to ACSF, presumably reflecting the combined kinetics of passive diffusion, active clearance and degradation mechanisms. In all experiments, birds were prevented from singing for at least 1.5 h after being switched from APV to ACSF to provide time for APV washout. For quantification of learning expressed immediately after training (Fig. 3f), we analysed the first songs performed after this

period. To further ensure that persisting effects of APV would not cause an underestimation of learning in our primary representations of the data (Fig. 3d, e), expression of learning was assessed the morning after the training period. This allowed sufficient time for the APV-dependent block of AFP output to subside while providing limited opportunity for the birds to sing in the absence of white noise; this was important because, in control conditions, singing in the absence of white noise results in a gradual loss of learned changes to fundamental frequency⁵ (that is, extinction). In a subset of experiments (8 of 24), white noise training was terminated (and APV was switched to ACSF) at least 3 h before sleep. In these experiments we found that the expression of learning before sleep was significantly greater than zero ($0.95 \pm 0.25\%$ change in FF, $P < 0.02$, signed-rank test) and only slightly less than learning the next morning ($1.3\% \pm 0.18\%$ change in FF). This indicates that washout of APV, independently of a period of sleep, was sufficient to enable the expression of learning. Probe position in RA was established by using electrophysiological mapping of RA during implantation and confirmed *post mortem* by identifying cannula tracts in brain sections stained for Nissl bodies. Additionally, in three birds, biotinylated muscimol (diluted to 500 μM ; EZ-link biotin kit; Pierce) was dialysed across the diffusion membrane to estimate the path of diffusion from the membrane¹⁴. In these birds, probe position was determined *post mortem* by histological staining for biotin and by comparing interleaved sections stained for Nissl bodies. Spread of drug outside RA tended to be in regions dorsal to RA, along the cannula, but not into the lateral areas where nucleus Ad is located.

Experiments with reversible inactivation of LMAN by reverse microdialysis.

We examined the progression of learning for data from experiments in which we transiently inactivated LMAN by using the same reverse dialysis technique that we used for infusing APV into RA¹⁴. To inactivate LMAN, we switched the dialysis solution from ACSF to the GABA_A agonist muscimol (100–500 μM ; Sigma; 3 birds, 12 experiments) or the Na⁺ channel blocker lidocaine (2%; Hospira; 1 bird, 2 experiments) at a flow rate of $1 \mu\text{l min}^{-1}$. Inactivations lasted for 3–4 h, during which a $1 \mu\text{l min}^{-1}$ flow rate was maintained. At the conclusion of inactivation, the dialysing solution was switched back to ACSF. We applied white noise contingent on FF over a total period of 2 days or more, during both control and LMAN inactivation periods. The threshold for escaping white noise was raised incrementally to drive progressive changes in FF. In each experiment, FF eventually reached a stable value because we stopped raising the threshold. We only considered LMAN inactivations on days before FF reached this stable value, to ensure that the bird retained the capacity for further learning. For each LMAN inactivation, learning after training was quantified as the difference in FF between the last 50 renditions of the syllable before infusion of drug and the first 50 renditions of the syllable after drug washout, normalized as for experiments with APV in RA. We excluded the first 1 h after switching the infusion solution to ACSF to permit washout. During the period with LMAN inactivated, which lasted a minimum of 3 h, the threshold for escaping white noise was set so that more than 10% but less than 50% of syllables escaped and thus a learning signal of differential reinforcement was present in each experiment. This is crucial for interpretation of the lack of learning in these experiments, because learning in this model does not proceed without such differential reinforcement⁶. Learning during training with LMAN inactivated was quantified with a linear regression of FF on the renditions of the targeted syllable during training with LMAN inactivated. For each inactivation, matched learning in control conditions was quantified by calculating the average hourly rate of change in FF during ACSF infusion on the day of that inactivation and multiplying that rate by the number of hours for which LMAN was inactivated. Probe positioning and the path of drug diffusion were evaluated *post mortem* by histological staining of sectioned tissue as described previously¹⁴. Tissue damage caused by cannulae enabled confirmation that probes were accurately targeted to LMAN. In addition, biotinylated muscimol or ibotenic acid was used to estimate the spread of diffusion, as described previously¹⁴.

Analysis. All analyses were performed with custom software written in MATLAB (Mathworks). For a given syllable, FF was measured over a consistent time window aligned to syllable onset; for syllables targeted with white noise feedback, the measurement time window was centred on the median time at which feedback was delivered. FF was calculated as described previously⁶ for both targeted syllables and non-targeted syllables of the same song. Spectral entropy, volume and duration were calculated as described previously⁵. Statistical significance was tested with non-parametric statistical tests; Wilcoxon signed-rank tests and Wilcoxon rank-sum tests were used where appropriate.