nature neuroscience

© 2006 Nature Publishing Group http://www.nature.com/natureneuroscience

Selective enhancement of associative learning by microstimulation of the anterior caudate

Ziv M Williams & Emad N Eskandar

Primates have the remarkable ability to rapidly adjust or modify associations between visual cues and specific motor responses. Whereas little is known as to how such adjustments in behavioral policy are implemented, recent learning models suggest that the anterior striatum is optimally positioned to have a role in this process. We recorded from single units and delivered microstimulation in the striatum of rhesus monkeys performing an associative learning task. Caudate activity during reinforcement was closely correlated with the rate of learning and peaked during the steepest portion of the learning curve when new associations were being acquired. Moreover, delivering microstimulation in the caudate during the reinforcement period significantly increased the rate of learning without altering the monkeys' ultimate performance. These findings suggest that the caudate is responsible for implementing selective adjustments to the 'associative weights' between sensory cues and motor responses during learning, thus enhancing the likelihood of selecting profitable actions.

Reinforcement-based associative learning is a critical adaptive mechanism that allows animals to effectively associate or 'link' between specific sensory cues and motor responses that are likely to lead to reward. Although the neuronal mechanisms underlying the acquisition of selective visual-motor associations remain unclear, there is increasing anatomic and physiologic evidence to suggest that the anterior striatum is key to this process. Medium spiny neurons in the anterior striatum receive projections from associative cortical areas presumed to have a role in learning and motor selection and project back to those same areas by way of the globus pallidus and thalamus[1–6]. In addition, it is likely that the same neurons are modulated by phasic dopamine release from neurons in the ventral tegmental area and the substantia nigra pars compacta (SNpc) that are sensitive to actual and predicted reward[4,7]. Experiments using intracranial self-stimulation (ICSS), in which rodents repetitively perform actions to receive electrical stimulation in dopamine-rich areas of the midbrain and striatum, have provided an especially powerful tool for examining reinforcement-based learning[8,9]. Although these experiments principally rely on the performance of single stereotypical actions, they indicate a possible anatomic basis for the enhancement of 'rewarded' behavior. Experiments using *in vivo* preparations also demonstrate that ICSS of the SNpc can lead to potentiation of corticostriatal synapses that are behaviorally relevant[10], further suggesting that the formation of specific associations probably depends on convergent activation from associative cortical areas and dopaminergic areas sensitive to reward.

On the basis of these observations, recent learning models have hypothesized that the striatum may provide a key function in learning by adjusting or modifying the 'associative weights' between sensory cues and motor responses that are likely to lead to reward[4,9–13]. In practical terms, these models suggest that when a particular pairing of a sensory cue and a motor response is followed by reward, the striatum acts to enhance their association, thereby increasing the animals' likelihood of selecting the same action on subsequent cue iterations.

In the current study, we recorded single-unit activity and delivered electrical microstimulation in the anterior striatum of awake behaving primates trained to concurrently learn discrete new visual-motor associations. We found that neuronal activity recorded in the caudate was closely correlated with the change in the monkeys' behavior during learning. Moreover, delivering microstimulation in the caudate during the reinforcement period significantly enhanced the rate of learning for specific associations without altering their ultimate acquisition or maintenance. Stimulation did not alter performance for concurrently learned nonstimulated associations and had little effect on learning performance when delivered in the putamen.

RESULTS

Learning behavior

We trained two rhesus monkeys to perform an associative learning task. During the task, the monkeys learned, by trial and error, to associate new visual images with specific joystick movements in one of four radial directions (Methods). In a given block of trials, we used four images. Two were randomly selected from a group of highly familiar images, and two were new images that the monkey had not seen before. A feedback tone indicated whether or not the correct target was reached, which was then followed by the actual reward (**Fig. 1a**). We presented an average of 16.5 ± 0.6 (mean ± s.e.m.) new images for each recorded cell over numerous trial blocks, such that multiple images and associated target locations were tested for each cell.

MGH-HMS Center for Nervous System Repair, Department of Neurosurgery, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. Correspondence should be addressed to E.N.E. (eeskandar@partners.org).

Received 23 December 2005; accepted 7 February 2006; published online 26 February 2006; doi:10.1038/nn1662

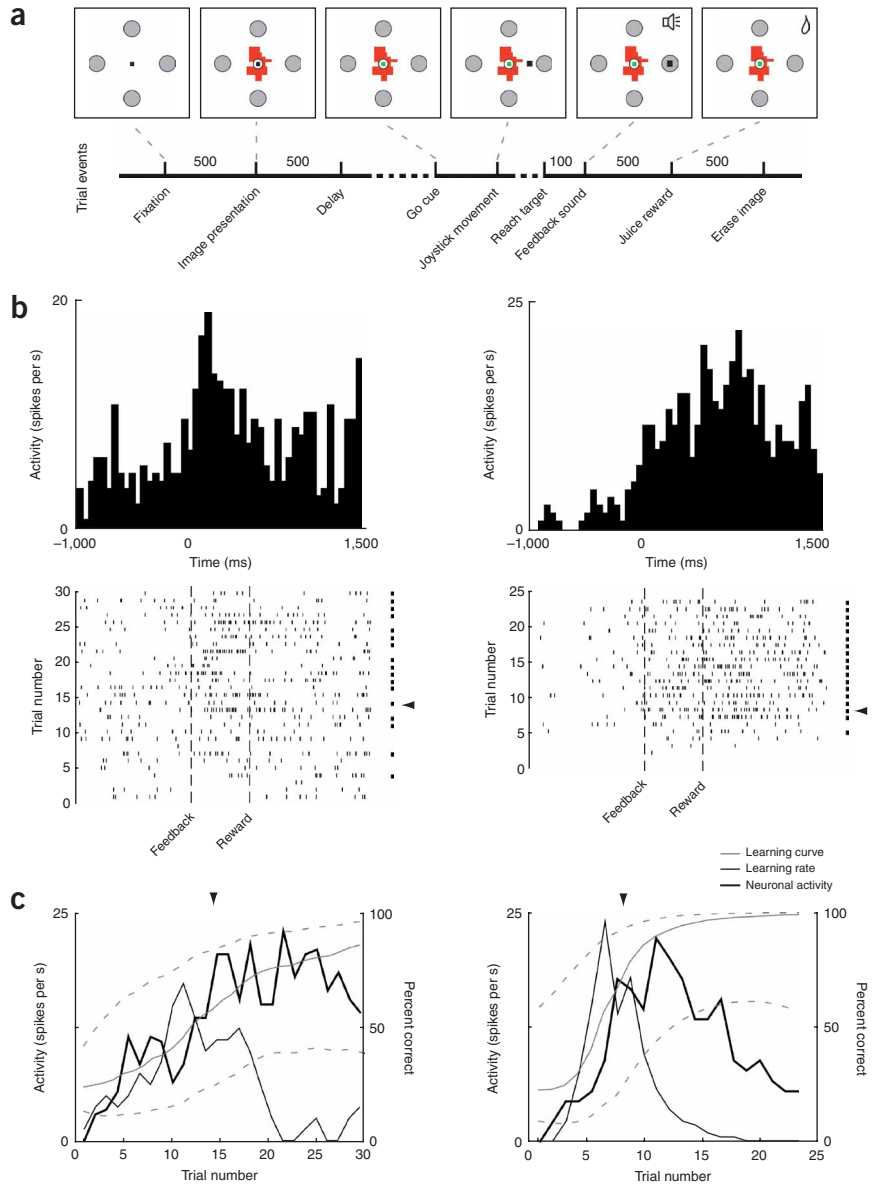

Figure 1 Single neuronal responses during learning. (**a**) Schematic illustration and timeline of events for the main task. (**b**) Perievent histogram and rasters for two single neurons over the course of one learning block while the monkey learned to associate a new image with a movement toward a specific target location. Activity is aligned to the feedback tone (first dashed line). Each row in the raster represents a single trial; the trial numbers are shown on the left. Black squares at right, correct trials; arrowheads, trial at which the learning criterion was reached. (**c**) Learning performance and mean neuronal activity during the feedback period. Gray line, monkeys' behavioral performance (learning curve); dashed lines, upper and lower confidence bounds (99%) estimated from the monkeys' performance. Black line, learning rate (first derivative or slope of the learning curve). Thick black line, average firing rate for the cell during the feedback period. In the left plot, neuronal activity closely correlates with the learning curve, whereas in the right plot, neuronal activity correlates more closely with the learning rate. Arrowheads, trial at which learning criterion was reached.

© 2006 Nature Publishing Group http://www.nature.com/natureneuroscience

Of these, the monkeys learned 12.4 ± 0.5 to criterion. The learning criterion was defined by the state-space model and was reached by the monkeys after an average of 4.1 ± 0.2 consecutive correct trials[14,15]. The monkeys reached criterion after an average of 9.7 ± 0.1 total trials (correct and incorrect).

Neither monkey demonstrated a response bias to a particular direction of movement at the start of the trial block (χ^2 test, $P > 0.05$), and the monkeys rarely selected the same incorrect target two or more times in a row. Furthermore, we found no difference in the number of trials it took the monkeys to reach learning criterion between images that were presented first or second during each block (rank sum test, $P > 0.05$).

Neuronal responses during learning

We recorded 171 cells from the dorsal anterior caudate and 72 cells from the rostral putamen. Of these, 153 caudate and 64 putaminal cells were classified as phasically active neurons (PANs), probably corresponding to medium spiny projection cells[16]. Tonically active neurons were not included in the current analysis[17]. The PANs recorded from the caudate are the main focus of this report, although the putaminal neurons provide an important contrast and will be discussed in that context.

Among caudate neurons, 23% were modulated during image presentation, 37% during movement, 39% during feedback and 40% during reward delivery (repeated-measures analysis of variance (ANOVA), $P < 0.05$; **Fig. 1b** and **Table 1**). The most prominent learning-related responses were observed during the feedback and reward periods. In 28% of neurons, there was a gradual increase and in 24% a gradual decrease in activity that plateaued once learning occurred (bootstrap test, $P < 0.01$; **Fig. 1c**, left). However, in a larger proportion of neurons (41%), there was an increase in activity that peaked near the time when the learning criterion was reached, when associations were being acquired, and then gradually decreased once the associations were made (**Fig. 1c**, right).

Learning curve and learning rate related neuronal activity

Neuronal responses during the different periods of the task were examined as a function of the monkeys' trial-by-trial performance, the learning curve, and the slope of this function, the learning rate (Methods). The learning rate indicates the rate of change in behavior that occurs during learning. Thus, low rates indicate no change in behavior, whereas high rates indicate periods of rapid change in behavior. The learning rate peaks during the steepest portion of the learning curve, at which time the animals are actively acquiring new associations between new images and particular movements[4,13]. Once an association is learned, however, there are no further changes in behavior and the learning rate returns to a low level.

Across the population of caudate cells, the most robust learning-related changes in activity were found during the feedback period (ANOVA, $P < 0.0001$; **Fig. 2a**,**b**). Activity during feedback correlated

Table 1 Neuronal modulation in the caudate and putamen

	Image	Movement	Feedback	Reward	Any	Total
Caudate	35 (23%)	57 (37%)	60 (39%)	61 (40%)	110 (72%)	153 (100%)
Putamen	18 (28%)	32 (50%)	37 (58%)	28 (43%)	58 (90%)	64 (100%)

Each column indicates the number and percent of cells modulated during each time period. The fifth column indicates the number and percent of cells modulated during any (one or more) of the four periods.

strongly with the learning rate (correlation coefficient of the learning rate: $r_r = 0.28$; H_0: $r = 0$, t-test , $P < 10^{-7}$), and only weakly with the learning curve (correlation coefficient of the learning curve: $r_c = 0.06$, $P > 0.05$; **Fig. 2c,d**). This was associated with a larger proportion of cells demonstrating a positive correlation with the rate of learning (**Fig. 2e,f**). In comparison, learning-related activity during image presentation was not significantly correlated with the learning rate ($r_r = 0.05$, $P > 0.05$) but was weakly correlated with the learning curve ($r_c = 0.11$, $P < 0.05$; **Fig. 2a,b**). On average, neuronal criteria preceded behavioral criteria by 1.8 ± 0.05 trials (**Fig. 2f**, inset). Learning-related changes during feedback were not present for familiar images ($r_r = –0.10$, $r_c = –0.12$, $P > 0.05$) or for images in which the monkeys did not reach learning criterion ($r_r = 0.04$, $r_c = 0.06$, $P > 0.05$). Learning-related changes were also not present for incorrect trials ($r_r = –0.04$, $r_c = 0.04$, $P > 0.05$). Caudate cells thus tended to show the greatest activity at feedback during the steepest portion of the learning curve when new associations were being acquired and less activity at the beginning and end of the curve when there was little or no change in learning behavior. Cells with learning rate–related activity were more prevalent anteriorly within the head of the caudate (linear regression, $P < 0.01$; **Fig. 3**).

In contrast to the caudate, the activity of rostral putaminal neurons was more closely correlated with the learning curve rather than with the learning rate. Most neurons demonstrated a progressive increase (47%) or decrease (20%) in activity with learning, whereas a smaller proportion (12%) demonstrated activity that peaked during the steep portion of the learning curve. Across the population, there was a significant correlation with the learning curve during feedback ($r_c = 0.39$; $P < 0.00001$) and little correlation with the learning rate ($r_r = -0.08$; $P > 0.05$; **Fig. 4**).

Controls for association formation, novelty and reward

A control task was introduced in which we tested 43 caudate cells to examine whether the observed responses depended on the monkeys actively learning the appropriate visual-motor associations. In this task, the sequence of movement directions was 'copied' from the previous block of trials onto a separate set of new images, but the direction of movement was instructed by a color change in one of the targets[11]. Thus, the monkeys were not required to learn the appropriate motor responses to the new images (**Supplementary Methods**) but still performed the same sequence of movements and experienced an identical succession of correct and incorrect feedback tones and rewards. In this task, there was no significant correlation between neuronal activity and progression through the sequence of trials

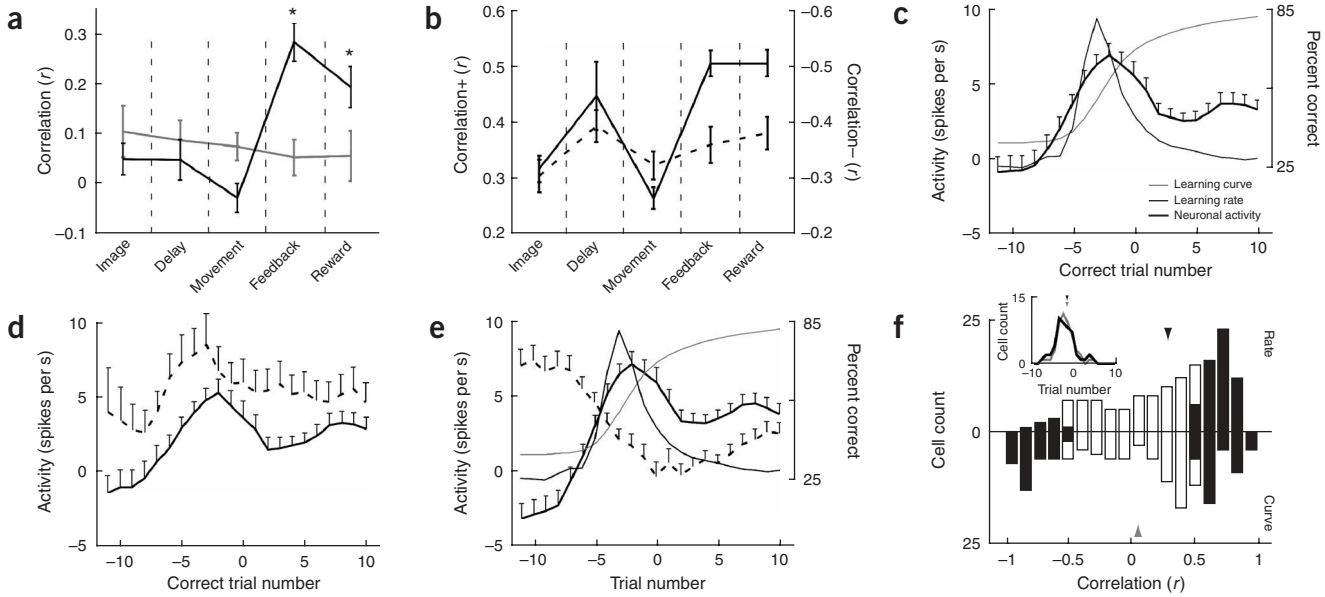

Figure 2 Population responses in the caudate during learning. (**a**) Correlation coefficients were calculated for the population of caudate cells ($n = 153$) by comparing neuronal activity during each 500-ms time period (that is, image, delay, feedback) with either the learning curve (gray line) or the learning rate (black line). Each point along the curve represents the mean r value for the population during that time period. Asterisks, intervals during which the distribution of r values was significantly different from chance ($P < 0.01$). (**b**) Positively correlated (solid line) and negatively correlated (dashed line) learning rate–related responses for the population during learning. (**c**) Mean neuronal activity and behavioral performance aligned to learning criterion (trial 0). Gray line, learning curve; black line, learning rate; averaged across all trials (confidence bounds not shown). Thick black line, mean neuronal activity during the feedback period minus baseline. Only cells with either a significant learning curve or significant learning rate–related activity were included in this plot ($n = 112$). (**d**) Mean neuronal activity during feedback for monkey P ($n = 90$, solid line), and monkey N ($n = 22$, dashed line). (**e**) Mean neuronal activity during feedback for positively (solid line) and negatively (dashed line) correlated responses. (**f**) Distribution of learning curve– and learning rate–related r values for all cells at the time of feedback. Black bars, cells with significant r values ($P < 0.01$). Gray and black arrows, mean learning curve–related and learning rate–related r values, respectively, for the population. Inset, distribution of lag times between neuronal criteria and behavioral criteria for learning-modulated cell. Gray and black arrows, mean lag for cells with significant learning curve–related and learning rate–related r values, respectively. Error bars indicate s.e.m.

© 2006 Nature Publishing Group http://www.nature.com/natureneuroscience

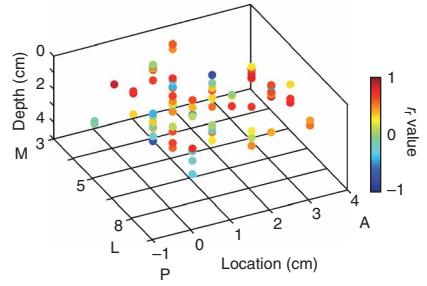

Figure 3 Locations of all recording sites within the caudate. Zero indicates the position of the anterior commissure in the anterior-posterior (A-P) dimension, the midline in the medial-lateral (M-L) dimension, and the dorsal margin of the caudate in depth. Color codes indicate the mean learning rate–related neuronal response during feedback (r_r) at each site.

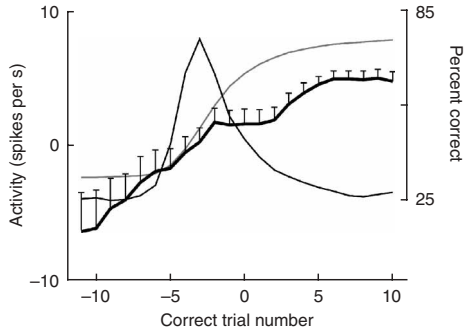

Figure 4 Population responses in the putamen during learning. Mean activity during feedback for putaminal cells ($n = 52$), with the same convention as in **Figure 2c**.

($r_r = -0.06$, $P > 0.05$; $r_c = -0.02$, $P > 0.05$), whereas the same cells showed significant learning rate–related activity in the standard task ($P < 0.05$). As in the standard task, there was little correlation between learning performance and neuronal activity on incorrect trials ($P < 0.05$).

Learning-related responses were not attributable to the use of unfamiliar stimuli. In another control task involving 53 cells, new images that had been learned in the previous block of trials were reintroduced in the following block but were now associated with an alternate set of target locations. In these trials, neuronal activity during feedback was again significantly correlated with the rate of learning ($r_r = 0.21$, $P < 0.01$; $r_c = 0.01$, $P > 0.05$).

Learning-related responses were also not due to simple changes in the frequency or schedule of reward delivery as learning progressed. In a control task involving 23 cells, we replayed an identical sequence of image presentations, cursor movements, feedback and reward delivery from the previous block of trials while the monkeys maintained fixation without moving the joystick. We found no peak in activity as the monkeys progressed through these trials ($r_r = 0.01$, $P > 0.05$) compared to the standard trials ($P < 0.05$). Moreover, we found no correlation between learning-related activity during feedback with arm movement within the peripheral target circles or eye movement within the fixation window (two-dimensional Kolmogorov-Smirnov test, $P > 0.05$). There was also no correlation between learning-related activity and licking activity following reward (bootstrap test, $P > 0.05$) suggesting that it was unlikely that this response pattern was due to simple changes in motor activity during the task.

Effect of microstimulation on learning

These findings suggested that neuronal activity in the caudate was closely correlated to the acquisition of new visual-motor associations when the change in learning behavior was the greatest. Nonetheless, these findings alone did not demonstrate that the caudate is causally involved in modifying the monkeys' actual behavior. To test this notion more directly, we used microstimulation[18,19] in the caudate and putamen. We hypothesized that if the caudate is indeed involved in modifying the association between sensory cues and motor responses during learning, then introducing microstimulation at the time of reinforcement may enhance or retard the rate at which specific associations are learned. We delivered high-frequency microstimulation (HFM) during the feedback-reward period following correct responses for one of the two concurrently learned new images in each block. As a control, these trial blocks were also interleaved with other blocks in which low-frequency microstimulation (LFM) or no stimulation (baseline) was delivered.

We found that performance on trials in which new images were coupled with HFM in the caudate was significantly better than performance on nonstimulated trials (two-tailed *t*-test, incremented

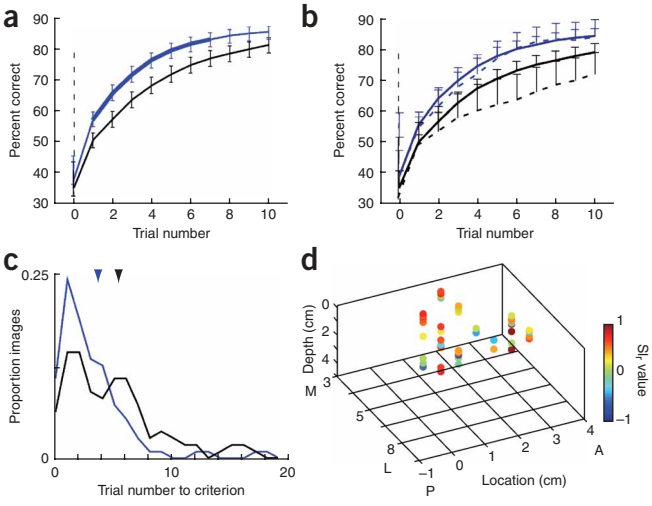

Figure 5 Effect of microstimulation in the caudate on learning. (**a**) Learning curves are aligned to the first correct trial (0). Blue line, mean performance for new images in which HFM was delivered in the caudate following correct responses during the reinforcement period (triggered at feedback). Black line, mean performance for new images in which no stimulation was delivered. Error bars indicate s.e.m (61 new images per curve). Thick areas along curve, trial points at which performance on stimulated trials was significantly different from performance on nonstimulated trials ($P < 0.05$). (**b**) Effect of microstimulation on learning for monkey P ($n = 51$ images, solid lines) and monkey N ($n = 10$ images, dashed lines), with the same convention as above. (**c**) Distribution of learning criteria for stimulated and nonstimulated images. Abscissa, total number of trials (beginning with the first correct trial) it took the monkeys to reach learning criterion. Arrowheads, average number of trials to reach criterion for all images in each set, with the same color convention as above. (**d**) Locations of all stimulation sites within the caudate, with location coordinates following the same convention as in **Figure 3**. Color codes indicate the mean SI_r for each site.

© 2006 Nature Publishing Group http://www.nature.com/natureneuroscience

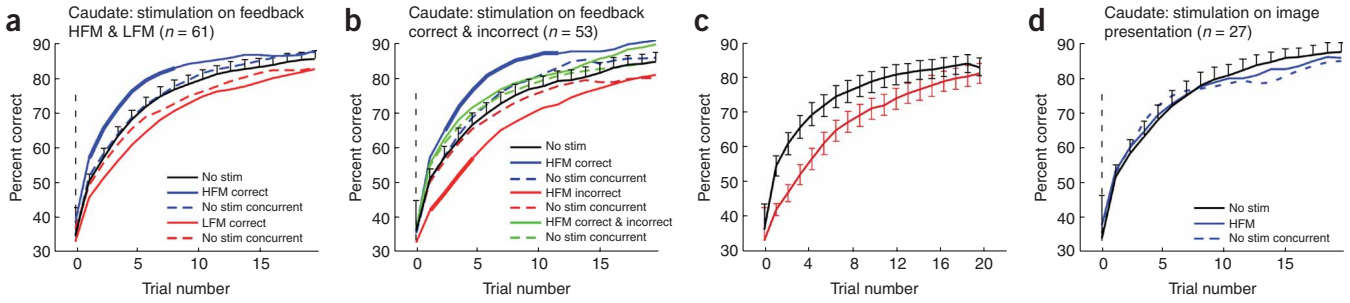

Figure 6 Stimulation controls. Learning curves are aligned to the first correct trial (0). Microstimulation trains were triggered either on feedback or on image presentation, depending on the experiment type (indicated in plot title). Solid colored lines in each figure represent the monkeys' performance for trials in which stimulation was delivered for one of the two concurrently learned new images. Dashed lines, with their corresponding color codes, represent the monkeys' performance for the concurrently learned nonstimulated new images. Black lines, trials in which no stimulation was delivered for either of the two new images (baseline). Error bars (s.e.m) are displayed only on the black curve for clarity. Thick areas along the curves, trial points at which performance on stimulated trials was significantly different from baseline ($P < 0.05$). The legend in the bottom right of each figure indicates the color code for each task type. The number of learning blocks used to construct each curve is indicated in parenthesis. All curves within each figure were obtained from interleaved trial sets performed within the same task sessions. (**a**) Comparison of LFM and HFM in the caudate. The black and blue curves shown here are the same as those shown in **Figure 5a**. (**b**) Stimulation in the caudate during correct trials, incorrect trials or both. (**c**) Comparison between baseline performance and performance on trials in which HFM was delivered for incorrect trials. The black and red lines shown here are the same as those shown in **b**. (**d**) Stimulation in the caudate during the image presentation epoch.

© 2006 Nature Publishing Group http://www.nature.com/natureneuroscience

comparisons, $P < 0.05$; **Fig. 5a**,**b**). This was associated with a significant increase in the steepness, or rate of rise, in the monkeys' learning performance (selectivity index of the rate of learning: $SI_r = 0.22$, *t*-test, $P < 0.001$). In addition, the monkeys took significantly fewer trials (30%) to reach learning criteria for stimulated new images compared to nonstimulated new images (rank sum test, $P < 0.05$; **Fig. 5c**). That is, the monkeys seemed to learn the correct visual-motor associations more quickly and reached learning criteria earlier when the correct responses were coupled with HFM during the reinforcement period. There was a slight tendency for microstimulation to exert a stronger effect on learning behavior anteriorly within the head of the caudate (linear regression, $P = 0.12$; **Fig. 5d**). In contrast to its effects on initial learning, HFM had little effect on the monkeys' final performance at the asymptote once the associations had been successfully learned (selectivity index of the final performance: $SI_{fp} = 0.01$, $P > 0.05$; **Fig. 6a**). No significant change in learning performance was found on LFM trials ($SI_r = –0.02$, $SI_{fp} = –0.01$; **Fig. 6a**).

Controls for response bias and motivation

Each image was associated with a unique movement direction. Hence, it was important to determine whether the improvement in learning could be attributed to a nonspecific response bias toward the stimulated target location. We therefore evaluated performance for nonstimulated new images that were learned concurrently with stimulated images but corresponded to a different direction of movement. If HFM led to a simple directional bias, then we would expect performance to deteriorate on concurrent nonstimulated trials. However, we found no difference in the rate of learning ($SI_r = 0.06$) or final performance ($SI_{fp} = 0.01$) on the nonstimulated trials compared to baseline (**Fig. 6a**). Similarly, there was no deterioration in performance on familiar trials, indicating that the effect of stimulation was selective for specific associations and did not arise from a nonspecific response bias.

It was also important to determine whether stimulation in the caudate was somehow perceived as being 'pleasurable'. In a separate task, monkeys were presented with a pair of targets that changed color simultaneously and from which the monkeys could select freely. Different paired combinations of reward and sitmulation were presented in separate blocks of trials. Upon reaching the target, the monkeys received either the standard reward, reward with HFM, HFM alone or none of the above. There was no preference in target selection for trials in which HFM versus no HFM was delivered or for trials in which reward with HFM versus reward alone was delivered (binomial test, $P > 0.05$; **Supplementary Fig. 1** online), suggesting that HFM did not lead to a simple hedonic response bias. The effect of caudate stimulation on learning therefore did not seem to lead the monkeys to favor a particular motor response but, rather, acted to enhance specific rewarded image-response associations.

Changes in learning performance with stimulation did not seem to result from a general increase in attention or motivational drive. When stimulation was delivered following both correct and incorrect responses, there was no difference in the learning rate ($SI_r = 0.06$) or final performance compared to baseline ($SI_{fp} = 0.02$; **Fig. 6b**). When stimulation was delivered on incorrect trials alone, the rate of rise in learning performance was significantly blunted ($SI_r = –0.25$, $P < 0.05$) and the number of trials to reach criterion was significantly prolonged (rank sum test, $P < 0.01$; **Fig. 6b**). Diminished performance was not due to repeated selection of trials in which HFM was delivered but, rather, due to a decreased likelihood of selecting the correct target location on subsequent trials (χ^2 test, $P > 0.05$). Final performance at the asymptote remained unchanged from baseline ($SI_{fp} = –0.05$; **Fig. 6c**). Moreover, when stimulation was delivered during the image

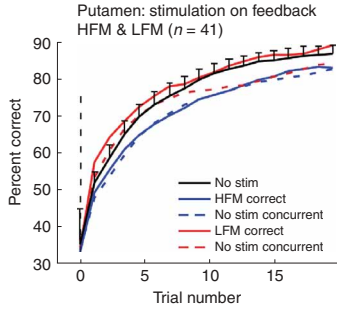

Figure 7 Effect of microstimulation in the putamen on learning. Same convention as in **Figure 6a**.

© 2006 Nature Publishing Group http://www.nature.com/natureneuroscience

presentation period in a separate set of trials, there was no significant effect on learning ($SI_r = –0.01$, $SI_{fp} = 0.02$; **Fig. 6d**), indicating that the delivery of stimulation in time periods other than the reinforcement period was not sufficient to alter learning performance.

In comparison to the caudate, there was no enhancement in learning when HFM was delivered in the rostral putamen. Rather, we noted a small nonsignificant decrease in performance ($SI_r = –0.16$, $P > 0.05$; $SI_{fp} = –0.01$, $P > 0.05$; **Fig. 7**), suggesting that the effect of stimulation in enhancing learning was selective to the caudate and was not generalized to other areas of the anterior striatum.

DISCUSSION

These findings suggest that the caudate is directly involved in augmenting selective visual-motor associations during learning and that this process probably occurs at the time of reinforcement. In practical terms, we propose the following simplified synthesis. The anterior caudate has convergent input from homologous associative cortical areas that respond to behaviorally relevant cues and from dopaminergic neurons that encode reward[1–7]. In the initial phases of associative learning, a new stimulus is coupled with some type of arbitrary motor behavior. By trial and error, a correct and thus rewarded response eventually occurs. This being an unexpected reward, there is a strong phasic burst of dopaminergic activity and the set of caudate neurons that is active becomes potentiated[7,9,10]. This process continues over the next few trials, reinforcing that particular visual-motor association. As the association is learned, the dopaminergic input is decreased, but the synaptic weights of that particular group of neurons remains enhanced as long as subsequent behavior remains coupled with reward.

Although the precise mechanisms by which microstimulation in the striatum exerts its effects are unknown, its influence on learning may be analogous to that of the dopaminergic responses observed with reward. For example, delivering tetanic electrical stimulation in the substantia nigra can lead to the long-term potentiation of activated corticostriatal synapses that are behaviorally relevant[10]. Additionally, in striatal slice preparations, long-term potentiation or depression of specific corticostriatal synapses is based on the combination of afferent cortical and dopaminergic activity[9,20]. By enhancing dopamine release[21,22], high-frequency microstimulation in the caudate may similarly result in the potentiation of particular corticostriatal synapses, or potentiate activated synapses directly.

The pattern of responses observed in the current study closely mirrors the actor-critic architecture proposed to have a role in associative learning. On the basis of this model, dopamine serves as the critic by signaling errors in predicted reward. The actor, in turn, uses the signal provided by the critic to implement direct changes to the monkeys' behavior in the form of a policy[4,11–13]. When unexpected reward is delivered shortly after a sensory cue is paired with a particular action, the actor implements changes to the 'associative weights' that link the specific stimulus to that particular action. Consistent with this model, we observed that caudate activity in single-unit recordings peaked when the change in behavior was greatest—at which time the strengthening of associative weights is hypothesized to occur[4]. Once the associations were learned, however, there was little further change in behavior, and neuronal activity again returned to a low level. In agreement with these findings, we observed that microstimulation in the caudate during reinforcement led to a significant increase in the slope of learning, or the rate of change in the monkeys' behavior. This occurred without affecting the monkeys' ultimate performance once the associations were already learned, even though microstimulation was present for all rewarded responses. Furthermore, the effect was highly selective for actively represented visual-motor associations without altering the rate of learning for concurrently learned new or highly established familiar associations. That is, only microstimulation that was temporally linked with a rewarded sensory-motor pair led to the selective augmentation of that particular association. The anterior caudate thus seemed to function as the actor, implementing direct changes to the monkeys' behavior in response to selective sensory cues.

The notion that the striatum may be involved in associative learning is supported by human imaging studies demonstrating caudate activation in subjects performing an instrumental conditioning task[11] and by studies in primates demonstrating activity changes correlated with learning[6,23–27]. What has been less clear, however, is whether the striatum is involved in implementing adjustments to the associations between sensory cues and the monkeys' actual motor behavior. Our findings suggest that the caudate acts to dynamically enhance or strengthen profitable behavioral policies during learning, thus increasing the organism's likelihood of selecting actions that lead to reward.

METHODS

Electrophysiology. Two adult male rhesus monkeys (*macaca mulatta*) were used in the study. A titanium head post, plastic recording chamber and scleral search coil were surgically implanted in strict accordance with guidelines set by the animal review committee at Massachusetts General Hospital. Neuronal activity was amplified, band-pass filtered between 200 Hz and 5 kHz, and sampled at 20 kHz. Spikes were stored and sorted offline using a template-matching algorithm (Spike 2, Cambridge Electronics Design). Eye position, joystick position and electromyographic licking activity were each sampled and recorded at 1 kHz.

Stimulation parameters. Biphasic stimulating pulses, with a cathodal phase leading, were delivered through a 300–600 kΩ tungsten microelectrode (Bak Electronics pulse generator and stimulus isolation unit)[19]. Phase length and interphase intervals were both 0.2 ms. Pulse frequency was either 20 Hz (low-frequency microstimulation) or 200 Hz (high-frequency microstimulation) at 200 μA. Stimulus trains lasted for 1000 ms and were triggered immediately after the feedback tone or image presentation depending on the experiment type. Stimulation epochs and durations were set to include either the reinforcement period (feedback and reward) or image presentation period (image presentation and delay).

Learning task. During the task, the monkeys were required to learn, by trial and error, to associate new visual images displayed at the center of the monitor with a specific movement. The monkeys used a joystick to guide a cursor emerging from the center of the screen to one of four targets displayed in the periphery. Each image was associated with only one rewarded target location and did not overlap with target locations associated with the other images. Two of the images were randomly selected from a group of familiar images that the monkeys were well-trained on, and the other two from a group of randomly generated new images that the monkeys had not previously seen. To ensure an even distribution of new and familiar trials, each of the four images randomly appeared once within a set of four consecutive trials. Trials in which the monkey selected an incorrect target location were repeated. After the monkeys completed a block of 18 correct trials for each image, a new set of images and target locations was selected. This learning process was repeated multiple times for each cell.

Behavioral analysis. The monkeys were considered to have successfully learned to associate a given new image with the correct target location if they had selected the correct target at least five times in a row. The probability of the monkeys having obtained this number of consecutive correct trials out of an average of 30 trials was 0.0192 and therefore unlikely to have occurred by chance. We used a state-space model to estimate the trial on which the monkey had first learned the association (that is, had reached learning criterion[14,15]). We defined the learning criterion as the first trial in which the lower 99% confidence bound obtained from the Gaussian state equation was greater than 0.25. The criterion describes the trial during which the monkey has first successfully made the correct association.

The mean behavioral performance, or learning curve, for each image was estimated from the monkeys' binary responses (correct or incorrect) using a Bernoulli probability model. This provided a continuous estimate, ranging from 0 to 1, of the monkeys' performance. The learning rate was calculated by approximating the central difference derivative of the learning curve, which describes the slope (that is, rate of change) in the monkeys' performance[28]. Parametric analysis of behavioral performance was obtained by fitting the learning curves to a standard logistic equation. Performance P_k at trial number $k = 1,\ldots,$ K, was defined by the equation

$$P_k = P_i + \frac{P_f - P_i}{1 + \exp(-\gamma(k - \delta))},$$

where P_i is defined as the initial performance and P_f is the final performance at the asymptotes. The rate of rise in the learning curve is governed by the constant γ, and the constant δ defines the inflection point of the curve. We found that 3% of the curves did not converge to the function using a Levenberg-Marquardt modification to the fitting algorithm; these were therefore excluded from further analysis[29]. There was no difference in mean residual values between the task types indicating that the goodness of fit was similar across trial conditions (one-way ANOVA, $P > 0.05$).

Selectivity indices (SI) for the learning rate (SI_r) and final performance (SI_{fp}) on stimulation trials were defined as (A – B)/(A + B), where A and B are the values for each of the two trial conditions (stimulated versus nonstimulated trials). The significance of change in SI was determined by using a one-tailed t-test (H_0: SI = 0, range –1 to 1; $P < 0.05$). Thus, a learning-rate SI (SI_r) of 0 would indicate that there is no difference in the rate of rise in performance (γ) between stimulated and nonstimulated trials. A positive SI_r would indicate that the learning rate was higher on stimulated trials, whereas a negative SI_r would indicate that the learning rate was lower.

Correlating neuronal activity with learning. Firing rates during image presentation, movement, feedback and reward time periods (500 ms) were correlated to the monkeys' learning performance by aligning neuronal activity for each of the new images to their corresponding learning criteria. Mean neuronal activity curves and behavioral performance curves were then obtained and the correlation coefficients (r) were calculated from the aligned data. This allowed us to evaluate the activity in each cell as a function of learning for multiple images and associated target locations before, during and after the criterion had been reached. Correct (rewarded) and incorrect trials were analysed separately. All cells were included for the analysis regardless of whether they demonstrated significant perievent activity. Cells with evidence of drift during baseline were not included at any point in the study.

We used a bootstrap test to evaluate the significance of correlation between neuronal activity and behavioral performance for individual cells. In each cell, neuronal activity was randomly shuffled 1,000 times across trials and correlation coefficients were recalculated from the shuffled data. If the rank of the calculated r value was higher than 99% or lower than 1% of the shuffled distribution, then the correlation was considered to be significant.

Learning criteria for neuronal responses were independently calculated for each cell by using a modification to the state-space model described previously[14,15]. Criterion for cells with significant learning rate–related responses were obtained by calculating the integral (area under the curve) of the response function. Criterion for cells with significant learning curve–related responses were obtained from the linear function of the response curve.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank E.N. Brown, W.F. Asaad, J.D. Macklis and R. Amirnovin for their thoughtful comments and critical review of the paper. Funding for this project was provided by the National Institutes of Neurological Disorders and Stroke, the Parkinson Disease Foundation and the Harvard Center for Neurodegeneration and Repair.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/natureneuroscience
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Alexander, G.E. Basal ganglia-thalamocortical circuits: their role in control of movements. *J. Clin. Neurophysiol.* **11**, 420–431 (1994).
2. Flaherty, A.W. & Graybiel, A.M. Input-output organization of the sensorimotor striatum in the squirrel monkey. *J. Neurosci.* **14**, 599–610 (1994).
3. Hoover, J.E. & Strick, P.L. Multiple output channels in the basal ganglia. *Science* **259**, 819–821 (1993).
4. Houk, J.C., Davis, J.L. & Beiser, D.G. eds. *Models of Information Processing in the Basal Ganglia* (MIT Press, Cambridge, Massachusetts, 1995).
5. McClure, S.M., Laibson, D.I., Loewenstein, G. & Cohen, J.D. Separate neural systems value immediate and delayed monetary rewards. *Science* **306**, 503–507 (2004).
6. Pasupathy, A. & Miller, E.K. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* **433**, 873–876 (2005).
7. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).
8. Fibiger, H.C., LePiane, F.G., Jakubovic, A. & Phillips, A.G. The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *J. Neurosci.* **7**, 3888–3896 (1987).
9. Wickens, J.R., Reynolds, J.N. & Hyland, B.I. Neural mechanisms of reward-related motor learning. *Curr. Opin. Neurobiol.* **13**, 685–690 (2003).
10. Reynolds, J.N., Hyland, B.I. & Wickens, J.R. A cellular mechanism of reward-related learning. *Nature* **413**, 67–70 (2001).
11. O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).
12. O'Doherty, J.P. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* **14**, 769–776 (2004).
13. 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).
14. Smith, A.C. *et al.* Dynamic analysis of learning in behavioral experiments. *J. Neurosci.* **24**, 447–461 (2004).
15. Wirth, S. *et al.* Single neurons in the monkey hippocampus and learning of new associations. *Science* **300**, 1578–1581 (2003).
16. Alexander, G.E. & DeLong, M.R. Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. *J. Neurophysiol.* **53**, 1417–1430 (1985).
17. Aosaki, T., Kimura, M. & Graybiel, A.M. Temporal and spatial characteristics of tonically active neurons of the primate's striatum. *J. Neurophysiol.* **73**, 1234–1252 (1995).
18. Cohen, M.R. & Newsome, W.T. What electrical microstimulation has revealed about the neural basis of cognition. *Curr. Opin. Neurobiol.* **14**, 169–177 (2004).
19. Tehovnik, E.J. Electrical stimulation of neural tissue to evoke behavioral responses. *J. Neurosci. Methods* **65**, 1–17 (1996).
20. Hernandez-Lopez, S., Bargas, J., Surmeier, D.J., Reyes, A. & Galarraga, E. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. *J. Neurosci.* **17**, 3334–3342 (1997).
21. Borland, L.M., Shi, G., Yang, H. & Michael, A.C. Voltammetric study of extracellular dopamine near microdialysis probes acutely implanted in the striatum of the anesthetized rat. *J. Neurosci. Methods* **146**, 149–158 (2005).
22. Cragg, S.J., Hille, C.J. & Greenfield, S.A. Dopamine release and uptake dynamics within nonhuman primate striatum *in vitro*. *J. Neurosci.* **20**, 8209–8217 (2000).
23. Hadj-Bouziane, F. & Boussaoud, D. Neuronal activity in the monkey striatum during conditional visuomotor learning. *Exp. Brain Res.* **153**, 190–196 (2003).
24. Hollerman, J.R., Tremblay, L. & Schultz, W. Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* **80**, 947–963 (1998).
25. Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V. & Graybiel, A.M. Building neural representations of habits. *Science* **286**, 1745–1749 (1999).
26. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* **1**, 411–416 (1998).
27. Miyachi, S., Hikosaka, O., Miyashita, K., Karadi, Z. & Rand, M.K. Differential roles of monkey striatum in learning of sequential hand movement. *Exp. Brain Res.* **115**, 1–5 (1997).
28. Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. *Numerical Recipes in C++: The Art of Scientific Computing* 2nd edn. (Cambridge Univ. Press, Cambridge, UK, 2002).
29. Schraudolph, N.N. Fast curvature matrix-vector products for second-order gradient descent. *Neural Comput.* **14**, 1723–1738 (2002).

© 2006 Nature Publishing Group http://www.nature.com/natureneuroscience