Changes in the Firing Patterns in Neurons of the Sensorimotor Striatum during Learning: What Changes and Why.

By

Terra D. Barnes

B.S. Biology
University of Illinois Urbana-Champaign (2001)

SUBMITTED TO THE DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES IN PARTIAL FULFULLMENT OF THE REQUIRMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN NEUROSCIENCE
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

March 2010
[JUNE 2010]

Signature of Author: _____
Department of Brain and Cognitive Sciences
March 10, 2010

Certified by: _____
Ann M. Graybiel, Ph.D.
Institute Professor
Thesis Supervisor

Accepted by: _____
Earl K. Miller, Ph.D.
Picower Professor of Neuroscience
Chairman, Committee for Graduate Students

Changes in the Firing Patterns in Neurons of the Sensorimotor Striatum during Learning: What Changes and Why.

By

Terra D. Barnes

Submitted to the Department of Brain and Cognitive Sciences on February 1, 2010 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Neuroscience

ABSTRACT

The basal ganglia, and specifically the sensorimotor (dorsolateral) striatum, have been implicated in stimulus-response learning. Here, I analyze the role the striatum plays in learning. We recorded from neurons of the sensorimotor striatum as rats learn, are over-trained, are extinguished, and are re-trained on a discriminative T-maze task. In this T-maze a gate was lowered immediately after an auditory click and the rats were allowed to proceed down the long arm of the maze. Mid-run, one of two tones was played. Rats had to choose to turn down either the left or right arm of the T-maze based on which tone was played. We discovered that population neural activity becomes restructured during learning and overtraining to emphasize the beginning and end of each trial. We also created a short-term memory version of the T-maze task by moving the location of the tone cue in order to determine if this affects the strength of the restructuring seen in the firing patterns of the striatum as learning progressed. Lastly, we examined the relationship the training induced pattern had to learning the tone-turn association and to other things that changed systematically throughout learning, such as speed.

Thesis Supervisor: Ann Graybiel, Ph.D.
Title: Institute Professor

TABLE OF CONTENTS

Chapter One

Introduction

Acknowledgements

Chapter One was has been previously published in Nature. 2005 Oct 20; 437 (7062): 1158-61. It was written by Dr. Ann Graybiel. Data was gathered by Dr. Dan Hu and Terra Barnes. Data was analyzed by Dr .Yasuo Kubota, Dr. Dezhe Jin, Dr. Ann Graybiel and Terra Barnes.

The data for Chapter Two was gathered by Dr. Jian Bin Mao. The data were analyzed by Terra Barnes, Dr. Ann Graybiel, Dr. Anna Dreyer and Dr. Emery Brown, Dr. Dan Hu, Dr. Dan Gibson and Dr. Yasuo Kubota. The data from Chapter Three was gathered by Dr. Dan Hu, Mark Howe, and Terra Barnes. The data were analyzed by Terra Barnes. I would like to thank all the people from the Graybiel laboratory especially Ann Graybiel, Katy Thorn, Dan Gibson, Hu Dan, Yasuo Kubota, Mark Howe, Ken Amemori, Hisham Atallah, Kyle Smith, Gila Fakterman, Theresa Desrochers-Moore, Pat Harlen, Brandy Baker, and Henry Hall, the people on my committee; Matt Wilson, Chris Moore and Terrence Sejnowski. I would also like to thank Andrew Young, Robert Ajemian and Dezhe Jin.

Introduction

The basal ganglia have been implicated in a variety of neurological disorders including Parkinson's disease, Huntington's disease, Tourette's syndrome, schizophrenia, obsessive compulsive disorder, certain dsytonias, addiction, and others (Kish SJ 1988; Freeman, Cicchetti et al. 2000; Murer, Tseng et al. 2002; Saka and Graybiel 2003; Sato, Sumi-Ichinose et al. 2008). In a number of these disorders there are gross motor deficiencies. This, together with the fact that the basal ganglia are intimately connected to several motor areas, lead to the belief that the major role of the basal ganglia was that of motor control. However, there is a large body of data – both anatomical and physiological -- that implicates the basal ganglia, and more specifically the striatum, in learning as well.

Anatomically, the striatum serves as the primary input structure to the basal ganglia and receives input from virtually all of the cortex and much of the thalamus (Flaherty and Graybiel 1994). It also receives a dopaminergic input from the substantia nigra pars compacta, sometimes referred to as a "learning signal" (Schultz 1998). In some cases, these dopaminergic inputs synapse on the neck of spines receiving input from the cortex, acting as a gate for the glutamatergic signal for the cortex. Therefore, the reward information that the learning signal is thought to possess can potentially control the level of contribution from cortex to the input cells of the striatum (Reynolds, Hyland et al. 2001).

The structures and connections within the striatum also suggest a role in learning. Within the striatum, there are mu-opioid receptor rich sub-regions called striosomes that are highly connected to limbic parts of the brain. Striosomes are interspersed among the

mu-opioid receptor poor matrix, which is connected to sensory and motor areas. These sub-regions have strikingly different levels of neurotransmitters, as well as different inputs, and, therefore, may have different roles to play in learning and motor movements (Saka and Graybiel 2003; Canales 2005). At the same time, input from one motor or sensory cortical area sends divergent input to several different matrisomes -- yet another type of functional sub-region within the striatum. This input reconverges as it is sent to the globus pallidus (Flaherty and Graybiel 1994).

Another prominent feature of the cytoarchitecture of the striatum is that it has two distinct pathways; the direct and indirect pathway. The direct pathway originates in the striatal medium spiny neurons that have D1 type dopaminergic receptors. These neurons project directly to and have a net excitatory effect on the thalamic output nuclei (via the globus pallidus internal and the substantia nigra pars reticulata). By contrast, the indirect pathway originates in the striatal medium spiny neurons that have D2 type dopaminergic receptors. These neurons project to the globus pallidus external which then projects to the sub thalamic nuclei which in turns projects to the globus pallidus internal and the substantia nigra pars reticulate resulting in a net inhibitory effect on the thalamic output nuclei (Mink 1996).

Finally, there are data to suggest that there are several closed-loop circuits between the striatum-thalamus and cortex that allow for further specialty and ever more complex analysis (Kelly RM 2004). Often, the thalamic nuclei receiving projections from the basal ganglia send projections back into the same frontal cortical regions of origin resulting in multiple closed cortico-basal ganglia loops. There are many regionally distinct areas of the striatum. The connections of the sensorimotor (dorsolateral) striatum

is functionally and anatomically distinct from the associative (dorsomedial) striatum (Yin and Knowlton 2006; Corbit and Janak 2007). Indeed these regions of the striatum evolve into two distinct brain regions in the primate; the putamen and caudate nucleus respectively. With all of this information flowing into the striatum and its complex cytoarchitecture, the striatum is well situated to play a role in learning.

Physiologically, there is evidence from human pathology that the basal ganglia, and specifically the striatum, play a central role in certain types of learning. Parkinson's disease is categorized by neuronal cell death of dopaminergic neurons in the substantia nigra pars compacta which project to the striatum. Patients with Parkinson's disease show deficiencies in learning procedural and sequential tasks (Mink 1996; Graybiel 2000; Siegert 2006; Yin and Knowlton 2006; Grahn, Parkinson et al. 2009). For example, Parkinson's disease patients show deficiencies in learning the probabilistic classification task in which several cues are probabilistically associated with two different outcomes. Patients must choose one of two cues and learn through repeated experiences, which cue has a higher probability of reward. Because of the probabilistic associations, this task makes it difficult to memorize the relationship between the stimuli. Patients with medial temporal lesions can learn this task normally and are thought to use a type of procedural learning. Parkinson's disease patients are impaired in learning these associations (Knowlton, Mangels et al. 1996).

Lesion studies in animals also support the idea that the striatum plays a role in instrumental learning. Divac and Kornorski showed that lesions of the striatum (and other parts of the basal ganglia) do not impair skilled movements or sequences of complex behaviors per se, but rather impair the ability to perform these actions in

response to well-learned stimuli (Divac, Rosvold et al. 1967; Konorski 1967; Yin and Knowlton 2006). Packard and McGaugh showed that inactivation of the striatum lead to a deficiency in rats' ability to use a striatal-dependent response strategy to learn to retrieve chocolate from a consistent location on a plus maze (e.g. they always turn right). Well-trained animals prefer to use this striatal dependent response strategy (Packard and McGaugh 1996). Adams and Kesner et al. showed that lesions of the sensorimotor striatum impair performance on a discriminative task in which well-trained rats were required to find a reward in one of two places (either on top or underneath an apparatus) based on which tone was played (Adams, Kesner et al. 2001). A variety of other studies have also demonstrated that lesions of the striatum impair animals' abilities to learn instrumental tasks (McDonald and White 1994; Knowlton, Mangels et al. 1996; Adams, Kesner et al. 2001; White and McDonald 2002; Yin, Knowlton et al. 2004; Featherstone and McDonald 2005).

Electrophysiological studies have further elucidated the function the striatum plays in learning by identifying behavioral correlates of striatal neuron firing and tracking changes in firing throughout learning. Recording in the putamen, the primate equivalent of the sensorimotor striatum, Samejima et al. trained two primates to choose to turn a handle either left or right. The animals were rewarded with either a large or small amount of water. The association between turn direction and reward probability was changed during each of five blocks of training per day. They found that cells in the striatum did not just encode direction of movement. In fact, more than a third of cells encoded the value of one of the two actions (left turn or right turn) during a delay period before the

choice was made. Their results suggest action values represented in the striatum could guide action selection and learning in the basal ganglia (Samejima, Ueda et al. 2005).

Firing patterns in the striatum have been shown to change systematically as rats learn a variety of tasks (Jog, Kubota et al. 1999; Chang and Gold 2003; Kitabatake, Hikida et al. 2003; Wickens, Reynolds et al. 2003; Henry H. Yin 2004; Samejima, Ueda et al. 2005; Tang, Pawlak et al. 2007; Hori, Minamimoto et al. 2009; Tang, Root et al. 2009; Yin, Mulcare et al. 2009). Tang et al. trained rats to move their head in an upwards motion in order to receive a water reward. They found the majority of cells in the sensorimotor striatum decreased their firing rate as training progressed but that a smaller subset increased or maintained their firing rate. They conclude that early in training the striatum uses a large population of neurons to modulate the behavior, but with habit development, the striatum uses fewer neurons to modulate or maintain the habitual movement (Tang, Pawlak et al. 2007). In our T-maze task, there is also evidence that cells in the striatum systemically change their firing patterns during learning. Jog et al found that as learning progressed in a discriminative T-maze task the percentage of cells responding to the beginning and end of trials increased while the percentage of cells responding during the middle of trials decreased (Jog, Kubota et al. 1999).

Here, I attempt to further elucidate the role the striatum plays in learning with a series of related studies on a discriminative T-maze task. In our discriminative task, rats hear one of two tones. Each tone is associated with chocolate reward at one end arm of the maze. Through trial and error, rats must learn which way to turn in order to receive reward. As rats learn this task, the percentage of neurons responding to the beginning and of trials increases while the percentage of neurons responding to events in the middle

of the task decreases (Jog, Kubota et al. 1999). Note that this task has three characteristics that distinguish it from a simple stimulus response task. First, it takes rats several days, often weeks, to become proficient at this task, which is true for many types of procedural learning (Neal J. Cohen 1985). Second, when rats are extinguished (i.e. reward is withheld), they continue to perform for several days despite the lack of reward. This is true of habits once they are formed (Yin and Knowlton 2006). Third, there is no approach to the conditional stimulus (the tone cue) in this maze. The rats learn an arbitrary association between a tone (that originates from a single speaker mounted at the choice point of the T-maze) and an end arm of the maze. In contrast, many studies require the animal to approach the conditioned stimulus (e.g. approaching a light, speaker or smell) (Atallah, Lopez-Paniagua et al. 2007). The underlying circuitry for discriminative approach learning is not necessarily the same as that used to learn an arbitrary association.

In **Chapter 2**, we classified the changes in neuronal firing patterns in the striatum that occur as rats are trained on our standard T-maze task. We found that early in training cells fire throughout trials and that later in learning cells fire more at the beginning and end of trials (Barnes, Kubota et al. 2005). We suggest that there is a process analogous to an explore/exploit model of learning that occurs at a cellular level during learning in the sensorimotor striatum.

This accentuation at the beginning and end of trials is especially interesting because the tone that tells the animal which direction to turn to receive chocolate reward occurs in the middle of the trial, when striatal activity is low. We suggest that the striatum may be chunking the entire motor program marking the beginning and end of the motor portion

of the task. To test this chunking hypothesis, we performed an experiment detailed in **Chapter 3**, in which the associative tone was presented at the beginning of the task. This manipulation allowed the animals to access all of the information necessary for successful task performance before beginning their motor sequence, and we hypothesized that this ability to pre-plan may enable stronger chunking. Consistent with this idea, we found that the beginning and end annunciation is much stronger in rats trained on the early-tone version of the task as compared to those trained on the standard version (Barnes et .al, *in prep*).

We further investigated which aspect of learning is responsible for the training-induced changes found in the striatal firing patterns. In **Chapter 4**, we performed an in depth analysis of the relationship between speed and firing rate in the sensorimotor striatum. We also include a control experiment in which animals received the same exposure to the task, experienced the same tones as rats trained on the classic task, and learned to run in the maze at the same speed. However, reward delivery for these rats was yoked to the average performance of the animals in **Chapter 2**, thus both groups of rats received similar levels of reward throughout training. Importantly, unlike the rats trained on the standard task, the rats trained on the non-associative version were not required to learn the tone-turn associations to obtain reward. The non-associative animals do not show the same accentuation of the beginning and end of the task as training progressed, suggesting that learning the cue-response relationship is necessary for acquisition of patterned neural activity in the sensorimotor striatum.

References

Adams, S., R. P. Kesner, et al. (2001). "Role of the Medial and Lateral Caudate-Putamen in Mediating an Auditory Conditional Response Association." Neurobiology of Learning and Memory **76**(1): 106-116.

Atallah, H. E., D. Lopez-Paniagua, et al. (2007). "Separate neural substrates for skill learning and performance in the ventral and dorsal striatum." Nat Neurosci **10**(1): 126-131.

Canales, J. J. (2005). "Stimulant-induced adaptations in neostriatal matrix and striosome systems: Transiting from instrumental responding to habitual behavior in drug addiction." Neurobiology of Learning and Memory **83**(2): 93-103.

Chang, Q. and P. E. Gold (2003). "Switching Memory Systems during Learning: Changes in Patterns of Brain Acetylcholine Release in the Hippocampus and Striatum in Rats." J. Neurosci. **23**(7): 3001-3005.

Divac, I., H. E. Rosvold, et al. (1967). "Behavioral effects of selective ablation of the caudate nucleus." Journal of comparative and physiological psychology. **63**(2): 184-90.

Featherstone, R. E. and R. J. McDonald (2005). "Lesions of the dorsolateral striatum impair the acquisition of a simplified stimulus-response dependent conditional discrimination task." Neuroscience **136**(2): 387-395.

Flaherty, A. W. and A. M. Graybiel (1994). "Input-output organization of the sensorimotor striatum in the squirrel monkey." J. Neurosci. **14**(2): 599-610.

Freeman, T. B., F. Cicchetti, et al. (2000). "Transplanted fetal striatum in Huntington's disease: Phenotypic development and lack of pathology." Proceedings of the National Academy of Sciences of the United States of America **97**(25): 13877-13882.

Grahn, J. A., J. A. Parkinson, et al. (2009). "The role of the basal ganglia in learning and memory: Neuropsychological studies." Behavioural Brain Research **199**(1): 53-60.

Graybiel, A. M. (2000). "The basal ganglia." Current Biology **10**(14): R509-R511.

Henry H. Yin, B. J. K. B. W. B. (2004). "Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning." European Journal of Neuroscience **19**(1): 181-189.

Hori, Y., T. Minamimoto, et al. (2009). "Neuronal Encoding of Reward Value and Direction of Actions in the Primate Putamen." J Neurophysiol **102**(6): 3530-3543.

Jog, M., Y. Kubota, et al. (1999). "Building neural representations of habits." Science **286**: 1745-1749.

Jog, M. S., Y. Kubota, et al. (1999). "Building Neural Representations of Habits." Science **286**(5445): 1745-1749.

Kelly RM, S. P. (2004). "Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits." Prog Brain Res **143**: 449-59.

Kish SJ, S. K., Hornykiewicz O. (1988). "Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. Pathophysiologic and clinical implications." N Engl J Med. **318**(14): 876-80.

Kitabatake, Y., T. Hikida, et al. (2003). "Impairment of reward-related learning by cholinergic cell ablation in the striatum." Proceedings of the National Academy of Sciences of the United States of America **100**(13): 7965-7970.

Knowlton, B. J., J. A. Mangels, et al. (1996). "A Neostriatal Habit Learning System in Humans." Science **273**(5280): 1399-1402.

Konorski, J. (1967). Integrative activty of the brain. Chicago, University of Chicago Press.

McDonald, R. J. and N. M. White (1994). "Parallel information processing in the water maze: Evidence for independent memory systems involving dorsal striatum and hippocampus." Behavioral and Neural Biology **61**(3): 260-270.

Mink, J. W. (1996). "THE BASAL GANGLIA: FOCUSED SELECTION AND INHIBITION OF COMPETING MOTOR PROGRAMS." Progress in Neurobiology **50**(4): 381-425.

Murer, M. G., K. Y. Tseng, et al. (2002). "Brain Oscillations, Medium Spiny Neurons, and Dopamine." Cellular and Molecular Neurobiology **22**(5 - 6): 611-632.

Neal J. Cohen, H. E. B. S. D. S. C. (1985). "Different Memory Systems Underlying Acquisition of Procedural and Declarative Knowledge^a." Annals of the New York Academy of Sciences **444**(Memory Dysfunctions: An Integration of Animal and Human Research From Preclinical and Clinical Perspectives): 54-71.

Packard, M. G. and J. L. McGaugh (1996). "Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning." Neurobiology of Learning and Memory **65**(1): 65-72.

Reynolds, J. N. J., B. I. Hyland, et al. (2001). "A cellular mechanism of reward-related learning." Nature **413**(6851): 67-70.

Saka, E. and A. M. Graybiel (2003). "Pathophysiology of Tourette's syndrome: striatal pathways revisited." Brain and Development **25**(Supplement 1): S15-S19.

Samejima, K., Y. Ueda, et al. (2005). "Representation of Action-Specific Reward Values in the Striatum." Science **310**(5752): 1337-1340.

Sato, K., C. Sumi-Ichinose, et al. (2008). "Differential involvement of striosome and matrix dopamine systems in a transgenic model of dopa-responsive dystonia." Proceedings of the National Academy of Sciences **105**(34): 12551-12556.

Schultz, W. (1998). "Predictive Reward Signal of Dopamine Neurons." J Neurophysiol **80**(1): 1-27.

Siegert, R. J. T., Kathryn D.; Weatherall, Mark; Abernethy, David A. (2006). "Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis." Neuropsychology **20**(4): 490-495.

Tang, C., A. P. Pawlak, et al. (2007). "Changes in activity of the striatum during formation of a motor habit." European Journal of Neuroscience **25**(4): 1212-1227.

Tang, C. C., D. H. Root, et al. (2009). "Decreased Firing of Striatal Neurons Related to Licking during Acquisition and Overtraining of a Licking Task." J. Neurosci. **29**(44): 13952-13961.

White, N. M. and R. J. McDonald (2002). "Multiple Parallel Memory Systems in the Brain of the Rat." Neurobiology of Learning and Memory **77**(2): 125-184.

Wickens, J. R., J. N. J. Reynolds, et al. (2003). "Neural mechanisms of reward-related motor learning." Current Opinion in Neurobiology **13**(6): 685-690.

Yin, H. H. and B. J. Knowlton (2006). "The role of the basal ganglia in habit formation." Nat Rev Neurosci **7**(6): 464-476.

Yin, H. H., B. J. Knowlton, et al. (2004). "Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning." Eur. J. Neurosci. **19**: 181-189.

Chapter Two

Activity of Striatal Neurons Reflects Dynamic Encoding and Recoding of Procedural Memories

Terra Barnes,[1*] Yasuo Kubota,[1*] Dan Hu,[1] Dezhe Z. Jin[1,2] and Ann M. Graybiel[1]

[1]*Department of Brain and Cognitive Sciences*
and the McGovern Institute for Brain Research
Massachusetts Institute of Technology
45 Carleton Street, E25-618
Cambridge, Massachusetts 02139 USA

[2]*Department of Physics*
Pennsylvania State University
0104 Davey Laboratory
University Park, Pennsylvania 16802 USA

*These authors contributed equally to this work

This paper was published in Nature. 2005 Oct 20;437 (7062):1158-61.

Abstract

Learning to perform a behavioural procedure as a well-ingrained habit requires extensive repetition of the behavioural sequence, and learning not to perform such behaviours is notoriously difficult. Yet regaining a habit can occur quickly, with even one or a few exposures to cues previously triggering the behaviour[1-3]. To identify neural mechanisms that might underlie such learning dynamics, we made long-term recordings from multiple neurons in the sensorimotor striatum, a basal ganglia structure implicated in habit formation[4-8], as rats were successively trained

on a reward-based procedural task, given extinction training and then given reacquisition training. The spike activity of striatal output neurons, nodal points in cortico-basal ganglia circuits, changed dramatically across multiple dimensions during each of these phases of learning. First, new patterns of task-related ensemble firing successively formed, reversed and then re-emerged. Second, task-irrelevant firing was suppressed, then rebounded, and then was suppressed again. These changing spike activity patterns were highly correlated with changes in behavioural performance. We propose that these changes in task representation in cortico-basal ganglia circuits represent neural equivalents of the explore-exploit behaviour characteristic of habit learning[9].

Results

The ability to establish habits, procedures and stereotyped behaviours brings great biological advantages to active organisms, and much evidence suggests that cortico-basal ganglia loops are critical for such learning[4-8, 10, 11]. If this view were correct, changes in the activity of basal ganglia neurons should accompany changes in behaviour not only as habits and procedures are initially acquired, but also as they are changed in response to altered behavioural contexts. To test for such restructuring of basal ganglia activity, we recorded chronically with multiple tetrodes for up to 63 sessions from the sensorimotor striatum of rats undergoing consecutive acquisition, over-training, extinction and reacquisition training on a conditional T-maze task (**Fig. 1, Supplementary Fig. 1 and Table 1**). The rats navigated the T-maze and turned right or left in response to auditory cues indicating whether chocolate reward was at the left or right choice-arm of the maze (**Fig. 1c**). This task requires trial-and-error learning, in which initial "exploration" of the

environment over successive trials leads, with successful learning, to "exploitation", in which correct choices are consistently made[9]. Performance accuracy increased during acquisition and was at or near asymptote during over-training (**Fig. 1d**). Accuracy then steadily deteriorated during extinction training, when reward was reduced (n = 4) or withheld entirely (n = 3), but recovered rapidly during retraining after extinction. Running times similarly fell, rose and fell (**Fig. 1e and f**).

As these behavioural changes occurred, the spiking of striatal neurons became redistributed across task-time (**Fig. 2, Supplementary Figs. 2 and 3**). We focused on the spike activity of neurons classified as striatal projection neurons, which directly participate in cortico-basal ganglia loop processing[12] (**Fig. 1a, Supplementary Fig. 1 and Supplementary Methods**). At the start of acquisition training, the spike responses of the task-responsive projection neurons, as a group, occurred throughout the maze runs (**Fig. 2a**). By the time the learning criterion had been met, however, the strongest per-unit firing occurred near the start and end of the runs. This progressive concentration of spike activity continued during the over-training period, even though behavioural performance had reached near-asymptotic values. In addition, early activity advanced from the time of locomotion onset toward the waiting period after the warning cue, and late activity shifted from around goal-reaching to around the end of turning (**Fig. 2a and c, Supplementary Figs. 2 and 3**).

These acquired spiking patterns were largely reversed during the extinction period (**Fig. 2a**). Mid-trial firing increased, and the temporal shifts, particularly for the early activity, reversed. When the reacquisition period then was initiated by returning the reward at the end of each correct run, there was another sudden shift in the spike patterns,

producing reduced mid-trial firing and a temporal advance of start activity resembling that seen during initial acquisition (**Fig. 2a, Supplementary Figs. 2-5**).

To estimate the randomness of the population spike activity across the entire trial time, we calculated the entropy of the average per-neuron firing across learning stages (**Supplementary Methods**). The entropy fell sharply during acquisition, rebounded during extinction, and fell again during reacquisition (**Fig. 2e**), in the absence of significant changes in average per-trial firing rates (**Supplementary Fig. 6**). The changes in the spike patterns were highly correlated with the changes in behavioural accuracy (**Fig. 2f and g**).

Remarkably, we found equally striking lability in the spiking patterns of the striatal projection neurons that lacked detectable phasic peri-event activity during the task (**Fig. 2b**). Some of these non task-responsive neurons fired at low rates both in-task and out-of-task, and some fired more out-of-task than in-task (Fig. 2d). The in-task activity of these neurons dwindled during acquisition and then nearly ceased. Yet, on the first day of extinction, the average per-neuron firing of these neurons returned to pre-training levels and remained elevated. Then, when reacquisition training began, their activity declined sharply. These abrupt shifts were evident whether the activity of the neurons during the task was classified with respect to pre-trial baseline firing (**Fig. 2b**) or was classified relative to in-trial activity (**Supplementary Fig. 2**).

To determine whether the tuning of task-related responses changed during learning, we measured multiple parameters (e.g., height, width, peri-event peak timing) of the phasic spike responses detected by a slope threshold (**Supplementary Methods**). None of these was altered during learning. By contrast, we found large-scale changes in the

proportion of spikes per entire trial-run that occurred within phasic responses (**Figs. 3a and b**). This proportion tripled during acquisition, fell abruptly during extinction and abruptly rose again during reacquisition (**Fig. 3b**). The number of phasic responses also successively changed (**Fig. 3c**). Reinforcing these redistributions of spike activity, the proportions of task-responsive projection neurons responding to different task events also progressively emphasized[7], de-emphasized and re-emphasized the beginning and end of the task (**Fig. 3d, Supplementary Fig. 6**). Notably, the sharpening of phasic responses during acquisition held not only for those occurring in the early and late parts of the task-runs in which overall spiking increased, but also for responses occurring in the middle parts of the runs in which spike activity decreased (**Supplementary Fig. 6**). This result suggests that even when fewer neurons responded, some "expert" responders with sharpened responses developed in the striatum as the task was acquired. This property, too, was subject to reversal and reappearance during subsequent extinction and reacquisition training.

Both the increase in concentration of spikes within phasic peaks during acquisition and the redistribution of spikes across run-times had the effect of reducing the spread of spiking across trial performance time as the rats learned the task. We looked for, but did not observe, significant changes in the variability of firing rates within peri-event or phasic-response windows across learning. However, we found major changes in the entropy (**Fig. 2e**) and in the variance (**Supplementary Fig. 6**) of spiking activity across the entire maze runs. Changes in spike distribution within the time frame of the entire procedural performance thus represented the key modulation of spike variability that we detected during learning.

Together, our findings demonstrate that per-trial spike distributions, response tuning and task selectivity were dynamically reconfigured as the procedural behaviour was acquired, extinguished and reacquired. Composite neural activity scores based on these factors were highly correlated with both behavioural accuracy and running times, especially during acquisition and extinction (**Fig. 4, Supplementary Fig. 7 and Supplementary Methods**). Restructuring of the day by day neural activity patterns in the "fast learners" (n = 5) but not in the "slow learners" (n = 2) early during acquisition (**Supplementary Fig. 8**) favoured a primary correspondence between the time evolution of the neural restructuring and associative learning. The acquired patterns were detectable in both correct and incorrect trials (**Supplementary Fig. 9**), however, so that the ensemble patterns were not tied to individual trial performance.

It has been proposed that the basal ganglia promote variability in behaviour during trial and error learning (exploration) and serve to evaluate behavioural changes to promote acquisition of optimal behaviour (exploitation)[10, 11, 13, 14]. Our findings suggest that there may be a direct neural analogue to such explore-exploit behaviour in the firing patterns of projection neurons in the sensorimotor striatum. We demonstrate two fundamental changes in the spike activity of striatal projection neurons during procedural learning. First, there was a global modulation of the firing of projection neurons. Early in training, the spike activity of the task-responsive population was spread throughout task time, as though all task events were salient (neural exploration). Even neurons without detectable phasic task-responsive activity fired at low rates during the task. Then, with continued training, this widespread spiking of the task-responsive population diminished, and their spike activity became focused (neural exploitation). At the same time, the non

task-responsive population fell silent, further reducing the task-irrelevant firing of the total projection neuron population. These changes in firing thus altered the distribution of striatal output neuron firing across the actual time-frame of the behaviour to be learned (the entire task run time). The reversal of the acquired task-related patterning during extinction and its reappearance in reacquisition fits the idea of increased neural exploration in the new extinction context and then a return to neural exploitation in the reinstated original context during reacquisition[15-17]. The vivid modulation of the spiking of striatal projection neurons without detectable task-related activity also accords with this interpretation.

Second, in the exploitation phases of learning, ensemble firing at the start and end of the learned procedure strengthened, and sharply tuned responses of "expert neurons" appeared. These changes suggest that early in training many candidate neurons fired, but that, with training and presumably competitive selection[18-20], the neurons with sharply tuned responses appeared, and, as a population, were tuned preferentially to respond near the start and end of the entire procedure performed. Our experiments leave open the question of where within the sensorimotor cortico-basal ganglia loop such changes were initiated. Because we recorded from striatal projection neurons, however, our findings demonstrate that such learning-related changes in neuronal firing occur as part of cortico-basal ganglia loop processing. The learning-related reduction in firing during the middle of task-time could indicate that striatal activity during this time was no longer needed for task execution, but could reflect the marking of behavioural boundaries in the process of chunking of the entire task performance[5]. These changing patterns could, in turn, reflect ongoing reorganization of cortico-basal ganglia activity[20-23]. If so, the patterns could

reflect neural representations related to the ready release of the learned behaviours by the appropriate context[5].

Cortico-basal ganglia circuits likely act in determining, through reinforcement-based evaluation, which actions to enhance or diminish as learning proceeds[4-6, 10, 11, 19, 20, 24-30]. Viewed in the context of such selection functions, our findings suggest that dynamic neural representations in the striatum could adjust the encoded salience of task events and behavioural responses as habits are formed, lost and regained.

Methods

The spike activity of neurons in the sensorimotor striatum was recorded chronically during behavioural training on a conditional reward-based T-maze task for 24–63 daily sessions from seven rats in which seven tetrode headstage assemblies had been implanted. Recordings began on the first day that the rats received training (about 40 trials per day) on the task, and were continued through successive acquisition training (stages 1–5), over-training (stages 6–15), extinction (stages 1–6) and reacquisition training (stages 1–6; Fig. 1b, Supplementary Table 2 and Supplementary Methods). In this task, rats learned to run down the maze and to turn right or left as instructed by auditory cues in order to receive reward. Behavioural data were acquired by means of photobeams and a CCD camera. Neural data (32 kHz sampling) were collected by means of a Cheetah Data Acquisition System (Neuralynx Inc.). Well-isolated units accepted after cluster cutting were classified as striatal projection neurons or interneurons (Supplementary Fig. 1b–d). Behavioural and neural data were aligned by time stamps and

were analysed by in-house software. The properties of both task-responsive and non-task-responsive projection neurons were analysed. Task-related responses of putative projection neurons were identified with respect to activity during a pre-trial 500-ms baseline period (threshold: 2 s.d. above baseline mean) and used to define task-responsive and non-taskresponsive populations (Supplementary Methods). Unit data were analysed per neuron and per neuronal population across task events (Fig. 1c). To analyse population activity, normalized firing rates were averaged for each learning stage, and indices of spike firing patterns across learning stages were computed. The proportions of neurons with different task-related response types, the proportions of spikes that occurred within peri-event phasic responses per session, and trial-to-trial spike variability were also calculated, along with composite neural scores and measures of the entropy of neural firing. Changes in these measures were compared to changes in per cent correct performance and running times of the rats across stages of training.

Acknowledgements We thank H.F. Hall, P.A. Harlan and C. Thorn for their help. This work was funded by the National Institutes of Health.

References

1. James, W. in The Principles of Psychology 104-127 (Dover, New York, 1890).
2. Dickinson, A. Actions and habits: the development of behavioural autonomy. Phil. Trans. R. Soc. Lond. B 308, 67-78 (1985).
3. Pavlov, I. P. in Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex (GV Anrep, Ed. & Trans.) (Oxford University Press: Humphrey Milford, London, 1927).
4. Packard, M. G. & Knowlton, B. J. Learning and memory functions of the basal ganglia. Annu Rev Neurosci 25, 563-93 (2002).
5. Graybiel, A. M. The basal ganglia and chunking of action repertoires. Neurobiol Learn Mem 70, 119-136 (1998).
6. Poldrack, R. A. et al. Interactive memory systems in the human brain. Nature 414, 546-50 (2001).
7. Jog, M., Kubota, Y., Connolly, C. I., Hillegaart, V. & Graybiel, A. M. Building neural representations of habits. Science 286, 1745-1749 (1999).
8. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19, 181-9 (2004).
9. Sutton, R. S. & Barto, A. G. Reinforcement Learning: An Introduction (MIT Press, Cambridge, MA, 1998).
10. 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).

11. Kao, M. H., Doupe, A. J. & Brainard, M. S. Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature 433, 638-43 (2005).

12. Wilson, C. J. in The Synaptic Organization of the Brain (ed. Shepherd, G. M.) 361-413 (Oxford University Press, New York, 2004).

13. Doya, K. & Sejnowski, T. J. in Advances in Neural Information Processing Systems. Vol. 7 (eds. Tesauro, G., Touretzky, D. S. & Leen, T. K.) 101-108 (MIT Press, Cambridge, MA, 1995).

14. Doya, K. & Sejnowski, T. in The New Cognitive Neurosciences (ed. Gazzaniga, M. S.) 469-482 (MIT Press, Cambridge, MA, 2000).

15. Bouton, M. E. Context and behavioral processes in extinction. Learn Mem 11, 485-94 (2004).

16. Routtenberg, A. & Kim, H.-J. in Cholinergic-monoaminergic Interactions in the Brain (ed. Butcher, L. L.) 305-331 (Academic Press, New York, 1978).

17. Myers, K. M. & Davis, M. Behavioral and neural analysis of extinction. Neuron 36, 567-84 (2002).

18. Gurney, K., Prescott, T. J. & Redgrave, P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biol Cybern 84, 401-10 (2001).

19. Graybiel, A. M., Aosaki, T., Flaherty, A. W. & Kimura, M. The basal ganglia and adaptive motor control. Science 265, 1826-1831 (1994).

20. Djurfeldt, M., Ekeberg, Ö. & Graybiel, A. M. Cortex-basal ganglia interaction and attractor states. Neurocomputing 38-40, 573-579 (2001).

21. O'Reilly, R. C. & Munakata, Y. Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Stimulating the Brain (MIT Press, Cambridge, MA, 2000).

22. Houk, J. C. & Wise, S. P. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. Cereb Cortex 5, 95-110 (1995).

23. Doya, K. Metalearning and neuromodulation. Neural Netw 15, 495-506 (2002).

24. Montague, P. R., Hyman, S. E. & Cohen, J. D. Computational roles for dopamine in behavioural control. Nature 431, 760-7 (2004).

25. Tanaka, S. C. et al. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat Neurosci 7, 887-93 (2004).

26. Reynolds, J. N. J., Hyland, B. I. & Wickens, J. R. A cellular mechanism of reward-related learning. Nature 413, 67-70 (2001).

27. Mink, J. W. The basal ganglia: focused selection and inhibition of competing motor programs. Prog Neurobiol 50, 381-425 (1996).

28. McClure, S. M., Berns, G. S. & Montague, P. R. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339-46 (2003).

29. Barto, A. G. in Models of information processing in the basal ganglia (eds. Houk, J., Davis, J. & Beiser, D.) 215-232 (MIT Press, Cambridge, MA, 1995).

30. Dickinson, A. & B., B. Motivational control of goal-directed action. Anim Learn Behav 22, 1-18 (1994).

Figures

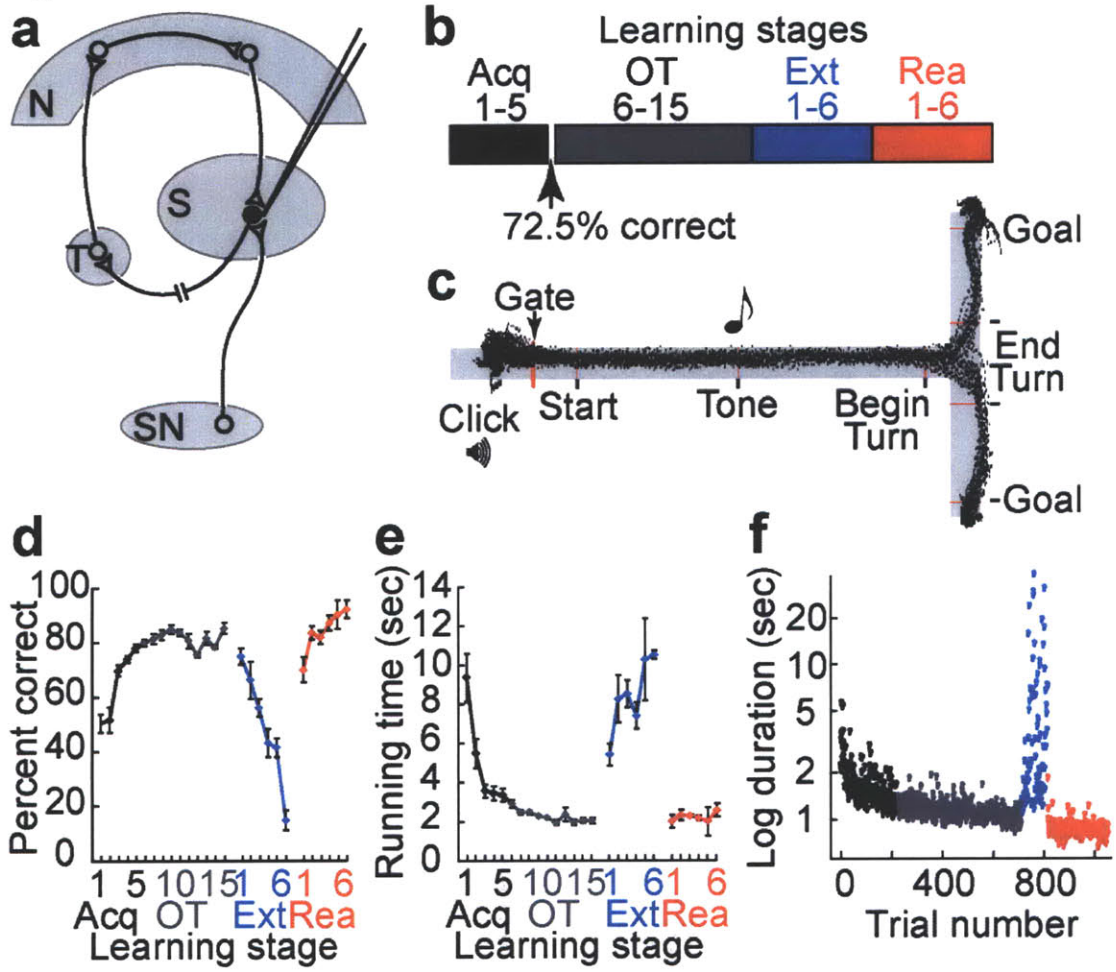


Figure 1 T-maze task and behavioural learning. **a**, Simplified cortico-basal ganglia circuit, indicating recording of striatal projection neurons. Neocortex (N), striatum (S), thalamus (T), substantia nigra (SN). **b**, Training stages (acquisition {}: 1-5, over-training [OT, grey]: 6-15, extinction [Ext, blue]: 1-6, reacquisition [Rea, red]: 1-6, described in Supplementary Methods). **c**, Run trajectories for one over-training session. **d-e,** Average percentages of correct responses (**d**) and average per-trial running times (**e**). Error bars represent standard errors of mean. **f**, Trial-by-trial running times for the tone onset to turn onset interval by a rat during successive training. Each dot represents performance in one trial.

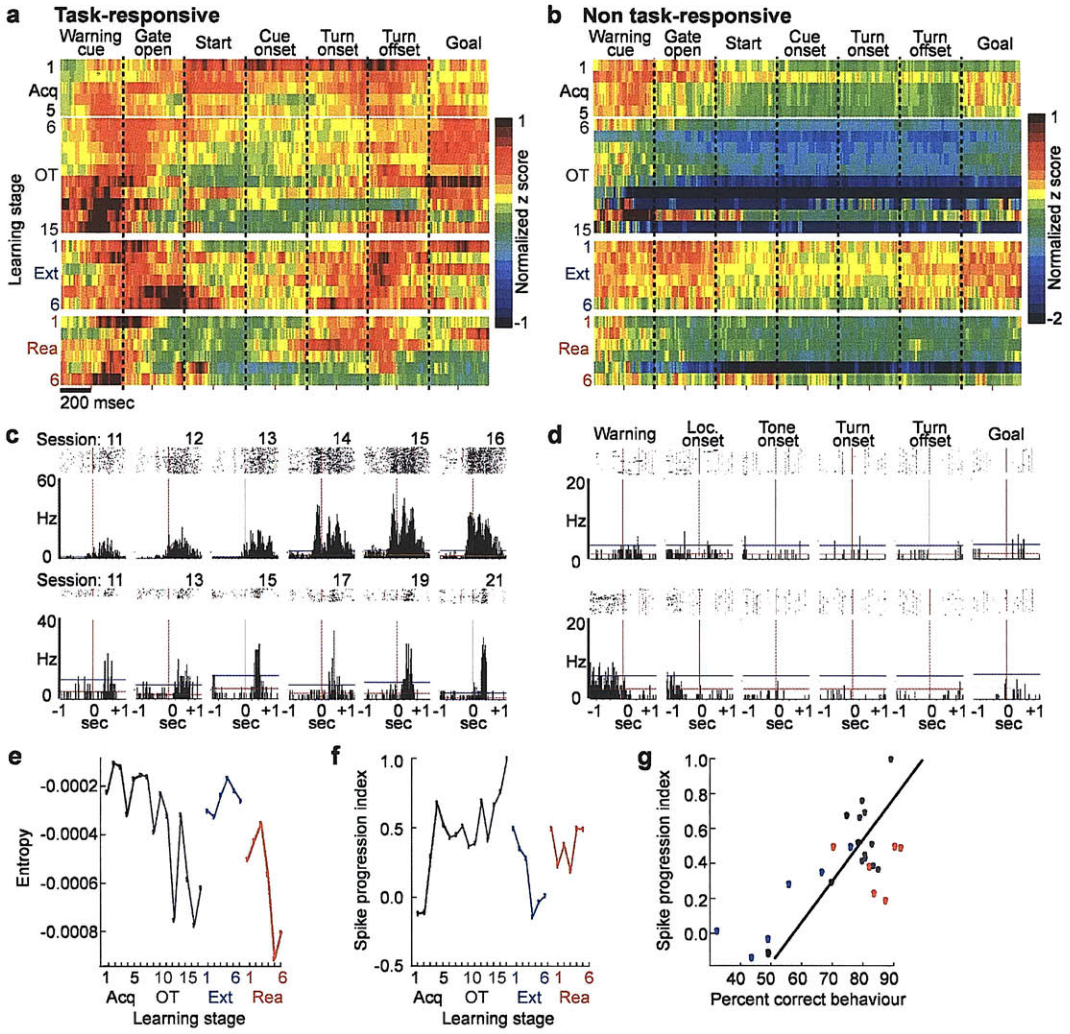


Figure 2 Plasticity in spike activity patterns of striatal projection neurons. **a-b**, Average activity of units classified as task-responsive (**a**) and non task-responsive (**b**) neurons plotted in 10-msec bins as z-scores normalized relative to that neuron's baseline activity, according to pseudo-colour scales shown at right, with one row per training session. Plots show ±200 msec time-windows around task events, abutted in the order of occurrence within a trial. **c**, Peri-event time histograms (PETHs, ±1 sec window) for units recorded on consecutive days at single sites (putative single units) illustrating strengthening and time-shift in responses around locomotion onset over 6 consecutive sessions (top) and sharpening of phasic responses at turn onset over 13 sessions (bottom). Horizontal lines indicate mean pre-trial baseline firing rates (red) and 2 SDs above the mean (blue, threshold for task-related activity). **d**, Typical PETHs for neurons lacking in-trial phasic peri-event activity ("non task-responsive" units). **e**, Entropy of per-trial spike activity of task-responsive units calculated for each training stage. **f**, Spike progression index (SPI) illustrating correlation of per-trial spike activity of task-responsive projection neurons at each training stage to the neural activity at the last stage of over-training. **g**, Significant correlation ($R = 0.74$, $P < 0.0001$) between SPI and progressive changes in percent correct behaviour during training. Acquisition and over-training (black), $R = 0.82$, $P = 0.0002$; extinction (blue), $R = 0.87$, $P = 0.02$; reacquisition (red), $R = 0.09$, $P = 0.87$.

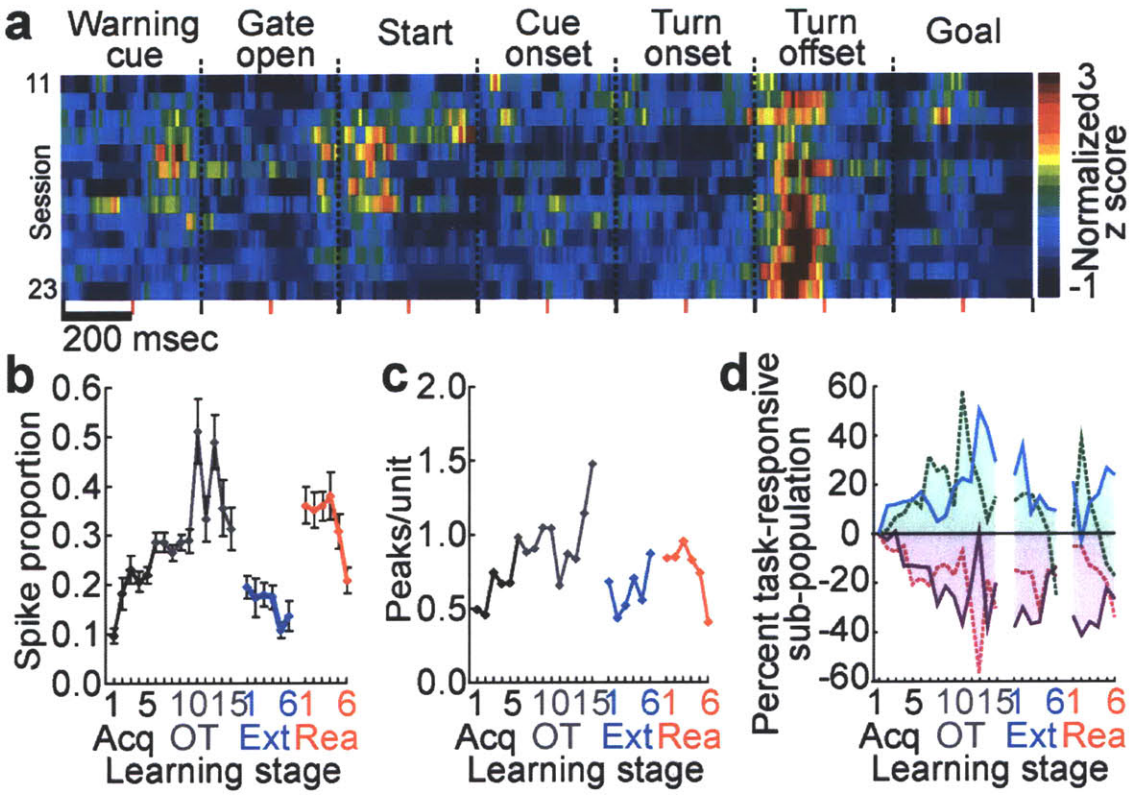


Figure 3 Multiple changes in projection neuron activity in the sensorimotor striatum during acquisition, extinction and reacquisition training. **a**, Unit activity at a single site recorded over 13 sessions. Each row represents a session. **b**, Proportions of spikes concentrated in phasic responses in each trial, averaged for each session. Error bars indicate standard errors of mean. **c**, Average numbers of phasic responses per unit. **d**, Percentages of task-responsive projection neurons with responses at warning cue (solid blue line), goal reaching (dotted green line), locomotion onset (solid purple line) or turning (dotted magenta line). Values plotted relative to first training stage.

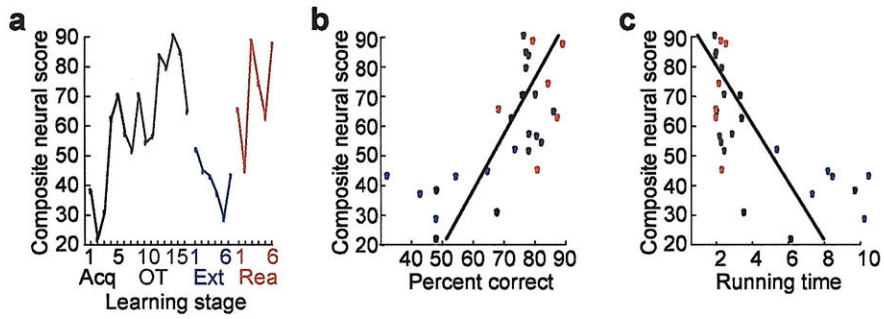


Figure 4 Striatal neural activity predictive of behavioural performance. **a**, Composite neural scores based on weighted neural measures at trial start (normalized per-neuron firing rates during the ±200-msec interval around warning cue, proportions of warning cue-responsive neurons, and proportions of spikes within phasic warning cue responses). **b**, Significant correlation between the composite scores (shown in **a**) and actual behavioural accuracy for each training stage (colour-coded as in **a**, $R = 0.69$, $P < 0.001$). **c**, Plot as in **b**, showing significant correlation between the composite neural scores and actual running times for each training stage ($R = -0.69$, $P < 0.001$).

Supplementary Methods

Experimental protocol. The spike activity of neurons in the sensorimotor striatum was recorded chronically during behavioural training on a conditional reward-based T-maze task for 24 to 63 daily sessions from seven rats in which seven tetrode headstage assemblies had been implanted. Recordings began on the first day that the rats received training (ca. 40 trials/day) on a conditional reward-based T-maze task and were continued through successive acquisition training (stages 1-5), over-training (stages 6-15), extinction (stages 1-6) and reacquisition training (stages 1-6, Fig. 1b, Supplementary Table 2). In this task, rats learned to run down the maze and to turn right or left as instructed by auditory cues in order to receive reward. Behavioural data were acquired by means of photobeams and a CCD camera. Neural data (32 kHz sampling) were collected by means of a Cheetah Data Acquisition System (Neuralynx Inc.). Well-isolated units accepted after cluster cutting were classified as striatal projection neurons or interneurons (Supplementary Fig. 1b-d). Behavioural and neural data were aligned by time stamps and were analyzed by in-house software. The properties of both task-responsive and non task-responsive projection neurons were analyzed. Task-related responses of putative projection neurons were identified with respect to activity during a pre-trial 500-msec baseline period (threshold: 2 SDs above baseline mean) and used to define task-responsive and non task-responsive populations. Unit data were analyzed per neuron and per neuronal population across task events (Fig. 1c). To analyze population activity, normalized firing rates were averaged for each learning stage, and indices of spike firing patterns across learning stages were computed, the proportions of neurons with different task-related response types, the proportions of spikes that occurred within peri-event

phasic responses per session, and trial-to-trial spike variability were also calculated, along with composite neural scores and measures of the entropy of neural firing, as described below. Changes in these measures were compared to changes in percent correct performance and running times of the rats across stages of training.

Surgical procedures. Headstages carrying tetrodes (200-250 KΩ) in each of seven independently-moveable microdrives (six for recording and one for reference) were mounted on the skull above an opening overlying the dorsolateral caudoputamen (AP = +0.5 mm, ML = 3.6 mm) in male Sprague-Dawley rats (250-350 g) anesthetized with ketamine (75-100 mg/kg) and xylazine (10-20 mg/kg). An anchor screw served as animal ground. All procedures were approved by the Committee on Animal Care of the Massachusetts Institute of Technology.

Behavioural procedures. Each rat was first handled in the animal colony room (3-5 days) and then was habituated to the T-maze chamber for 3-5 days. About one week after surgery, acquisition training began (Fig. 1b and c). In each trial, a warning cue (~70 dB click) was presented 250 msec before the opening of the start gate, while the rat was at the start location. When the gate opened, the rat was allowed to run down the maze. Half-way down the main alley, one of two tones (1 and 8 kHz pure tones, ~80 dB) was sounded to indicate which of the choice arm goals was baited with reward (chocolate sprinkles). The tones remained on until the rat reached one of the goals or the trial was terminated (Fig. 1c). Tone-goal arm assignments were randomized and counterbalanced

among rats. Approximately 40 trials separated by 1-3 minute inter-trial intervals were given each day.

Each rat was required to reach the correct goal in at least 72.5% of trials in a session to attain the acquisition criterion for significant correct performance ($p < 0.01$, chi-square tests) and then had to perform at or above this level in 10 out of 11 consecutive daily sessions to reach the over-training criterion. The numbers of initial acquisition training sessions ranged from 3 to 21, and the numbers of over-training sessions varied from 10 to 38. The rats were then given extinction training, in which reward was reduced to 1-3 trials per session ($n = 4$) or withdrawn altogether ($n = 3$). Extinction training lasted 2-11 days. Immediately thereafter, reacquisition training on the original task began and continued until the rat performed at the 72.5% correct criterion level or headstages failed (Supplementary Tables 1 and 2). Two to eleven reacquisition sessions were given. During all training phases, sessions were terminated if the rat stopped performing the task before completing 40 trials. Each day, recordings were made for an average of 38.1 trials during acquisition, 33.3 trials during extinction and 38.4 trials during reacquisition.

Neuronal and behavioural data acquisition. Tetrodes were gradually lowered through the brain toward the striatum (3.5-5.0 mm) during the 1-week recovery period after surgery. Once they reached the target, the position of each tetrode was adjusted until 3-5 distinguishable units appeared in the recordings. Task training then began. During training, tetrodes were moved as little as possible, and then in small (e.g., <100 μm) steps to maintain high quality, multiple single-unit recordings. The average distance of tetrode

movement throughout the recording periods is shown for each rat in Supplementary Table 1. We recorded an average of 10.8 units per daily session. The absolute numbers of units recorded could not be accurately determined, given probable repeated recording from individual neurons on successive days. Data were thus compiled in terms of units per session and were then averaged.

In selected sessions, sensory responses of recorded units were tested before or after behavioural training by tactile stimulation of contra-lateral body areas (e.g., front and hind limb, neck, back and body) with a glass stir-bar and by manipulation of joints. This examination identified sensory responses of units recorded by a tetrode, but did not provide information about which unit was activated by the stimulation. Despite this limitation, the results did not suggest any clear relationships between sensory responsiveness and task-related activity of recorded units.

Unit activity (gain: 200-10000, filter: 600-6000 Hz) was recorded during all training sessions with a Cheetah Data Acquisition System (Neuralynx Inc.). Spikes exceeding a preset voltage threshold were sampled at 32 kHz per channel and were stored with time stamps. The animal ground or a single tetrode channel served as reference. The movement of the rat was monitored continuously and recorded (sampling rate: 60 Hz) by a video tracker that received images from an overhead CCD camera. The times of occurrence of behavioural and stimulus events were determined either online by the use of photobeams (Med Associates, Inc.) or offline by analyzing the tracker data.

At the end of training, rats were deeply anesthetized (Nembutal, 50-100 mg/kg), and lesions were made to mark the final recording sites (25 μA, 10 sec). Rats were then perfused with 4% paraformaldehyde in 0.1 M phosphate buffer, and 30 μm thick

transverse frozen brain sections were stained for Nissl substance to identify recording tracks and lesion sites (Supplementary Fig. 1a).

Data Analysis.

1. Behavioural data. The performance of each rat in each training session was measured by the accuracy of responses (percent correct) and the time that elapsed as the rat ran the maze from gate opening to goal reaching (running time), averaged over all trials per session. Changes in these measures during training were analyzed by repeated measures analysis of variance (ANOVA). In order to combine data from different rats to detect learning-related changes in neural responses, we defined stages of learning according to the response accuracy in each training session as follows: stage 1 = first training session, stage 2 = second training session, stage 3 = first session with >60% correct responses, stage 4 = first session with >70% correct responses, and then subsequent stages as pairs of consecutive sessions with >72.5% correct performance. Some pairs of consecutive sessions were on consecutive days of training, but others were separated by gaps in which per-session performance fell below 72.5% correct (Supplementary Table 2). For extinction and reacquisition sessions, stages were: stage 1 = first training session, stage 2 = second training session, and then stages 3-6 = pairs of consecutive sessions.

2. Spike sorting and unit classification. Unit activity recorded by each tetrode was first sorted into single units by the use of AutoCut (DataWave Technologies) under manual control, and the quality of sorted units was tested by analyzing auto-correlograms and overlays of spike waveforms. Each unit was included for analysis if its total number of spikes exceeded a threshold of 100 spikes/session, and each accepted unit was classified,

as shown in Supplementary Fig 1b-d, as either a putative projection neuron, a putative fast-firing interneuron (FFN) or a putative tonically firing interneuron (TFN). Units classified as putative projection neurons made up 2091 of 3149 accepted units (66.4%). These were the focus of this study and, for convenience, are termed projection neurons in the text. Smaller numbers of units were classified as FFNs (n = 942, 29.9%) or TFNs (n = 116, 3.7%). The relatively small numbers of FFNs and TFNs precluded conclusive analysis of changes in their firing patterns across all task events and the 27 learning stages; but in the data available, we did not observe the large-scale, multiple changes in firing patterns that we found for the neurons classified as projection neurons.

3. Task-related activity of individual units. Peri-event time histograms (PETHs) were made for each unit for each time-stamped task event (warning cue, gate opening, locomotion onset, tone onset, turn onset and offset, and goal reaching). Task-related responses were defined as responses in which the spike counts in four or more consecutive 20-msec bins, with at least one of those bins occurring in ±200-msec peri-event time windows, had 2 or more spikes and exceeded the criterion level, which was set at two standard deviations (SDs) above the mean activity recorded during the pre-trial baseline 1900 to 1400 msec before warning cue. For units that did not fire during the baseline period, task-related responses were defined as epochs with four or more consecutive bins with spike counts of at least 2. The proportions of units with such event-related phasic discharges ("task-responsive units") were calculated for each task event for each learning stage. The remaining units were designated as "non task-responsive units." The proportions of task-responsive units increased from 55-80% of all accepted units during acquisition to 80-100% late in over-training, then decreased to 45-75% during

extinction and then rose during reacquisition to the levels found during acquisition to 60-80% of accepted units (Supplementary Fig. 6b).

4. Population firing profiles. Ensemble firing rates of projection neurons were calculated for consecutive 10-msec bins during the 500-msec pre-trial baseline period and for 2-sec time windows centered on each task event. The spike counts for each unit were first smoothed by taking running averages of three consecutive bins, and the smoothed spike rates were then converted to z-scores: $\frac{FR_i - FR_{mt}}{SD_{mt}}$, where FR_i is the smoothed firing rate in the i^{th} bin of the peri-event period, FR_{mt} is the mean firing rate over all peri-event periods, and SD_{mt} is the SD of firing rates for all peri-event periods, with values averaged for all trials of a session. For calculating z-scores, the mean and SD for all peri-event periods were used, rather than those during pre-trial baseline periods, because some units did not fire any spikes during the baseline period, preventing z-score calculation. Each z-score was then normalized to baseline by subtracting a z-score value corresponding to the mean baseline spike counts. These per-unit normalized z-scores were then averaged to construct peri-event spike histograms for groups of recorded units classified as putative projection neurons exhibiting particular task-related response profiles (Fig. 2, Supplementary Fig. 2). Increases and decreases in these average population responses were evaluated relative to the average baseline activity of the given neuronal population, defined as deviations from the baseline mean by over two SDs.

To calculate the randomness of the distribution of population spiking across the entire task time (maze runs), the z-score in the ith 10-msec bin within the ±200-msec peri-event windows was converted into the population firing rate f_i at the bin using $f_i = z_i + \alpha$, where α is a constant. We chose $\alpha = 1$, which made $f_i > 0$ for our data

set; the result of the randomness calculation was insensitive to the value of α. The firing rates were used to compute the probability density p_i of finding a spike in the bin with equation $p_i = f_i \Big/ \sum_{i=1}^{N} f_i$, where $N = 287$ is the number of the bins (41 10-msec bins for each of 7 task events). The entropy of the population spiking through the entire task was computed with $\text{Entropy} = -\sum_{i=1}^{N} p_i \log p_i + \log(1/N)$.

The strength of patterning in the population neuronal activity across the entire task time was measured by the structure index calculated for each training stage. The structure index was then defined as $\text{Structure Index} = \dfrac{\sum_{i=1}^{N} (z_i - \bar{z})^2}{N-1}$, where $\bar{z} = \dfrac{\sum_{i=1}^{N} z_i}{N}$ is the mean of the z-scores across all bins. The structure index was then compared across training stages.

Changes in population firing patterns across stages of training were evaluated by constructing correlation matrices between the z-score vectors representing population firing rates in the 10-msec bins within the peri-event windows at each learning stage. A spike progression index (SPI) was defined as the correlation of the z-score vector at each stage relative to the z-score vector at the last stage of over-training. These values were compared to changes in behavioural accuracy by computing Pearson's product moment correlations between the two data sets.

5. Measures of phasic responses. To determine whether the tuning of task-related responses changed during learning, the occurrence of phasic responses of each accepted unit in per-session PETHs was detected by a peak detection algorithm. Each PETH was first smoothed with a Savitzky-Golay filter with filter window size of 15 consecutive data

points. Noise fluctuations in the PETH were estimated by subtracting the smoothed PETH from the original PETH and calculating the standard deviation (SD) of the residuals. A noise detection threshold was set at 5 x SD. If the range of a given PETH (max - min) was less than the noise threshold, or the max was less than 1 Hz, the PETH for that unit was considered too noisy for analysis. Each accepted, smoothed PETH was analyzed by an in-house peak detection procedure for a curve y(x). At each data point, the slope of the curve was determined by least-square fitting with 21 neighboring data points. A slope threshold was set to be the range of y divided by the range of x, and if the slope at the data point exceeded this threshold, the data point was considered significant. Then, y values were scanned from small x to large x in order to find a set of 10 consecutive significant points with positive slopes and then a set of 10 consecutive points with negative slopes. A candidate peak was defined as lying between two such sets. The range of data points included in the candidate peak was determined by identifying the local minima to the left of the positive slope set and to the right of the negative slope set. Data points within the range of the candidate peak were least-square fitted to a Gaussian function with a linear base. The formula of the fit function is

$b_1 \exp\left(-b_2 \left(x - b_3\right)^2\right) + b_4 + b_5 x$, where $b_i, i = 1,..,5$ are fitting parameters. The x range of the peak was set to the range of the candidate peak, or to the half-width of the fitted Gaussian, whichever was smaller. Candidate peaks were rejected if the width was less than two data points or the width times the height was less than 0.1 times the range of x times the noise threshold.

These fitted functions were used to extract properties of the phasic responses such as absolute and normalized peak heights, peak widths at the base of the peak and at half-

height, peri-event response timing (timing of peak's maximum), and the proportion of spikes per trial that occurred within individual detected peaks. All of these measures were compared across learning stages by ANOVA. Some response peaks that were below the standard +2 SD criterial threshold for acceptance as phasic task-related responses were detected by the peak detection algorithm, which compared windowed firing to per-trial firing, not to baseline firing. This occurred in 16.6% of the units classified as non task-responsive units (15.0% during acquisition, 22.2% during extinction and 11.8% during reacquisition).

6. Trial-by-trial variability in spiking. Variability of spiking from trial to trial within a session was computed for each accepted unit for 200-msec intervals before, around and after each task event time stamp and for the durations of the phasic responses identified by the peak detection algorithm. The mean firing rates and their SDs were calculated based on per-trial spike counts within a peri-event or phasic response time window for each unit. Normalized trial-to-trial coefficients of variation (SD/mean firing rate) in spiking during the time window for subpopulations of units were compared across training stages by ANOVA.

7. Composite neural response measures. Composite neural activity scores were computed by a maximization procedure applied to neural data for normalized per-neuron firing rates, proportions of task-responsive neuronal subpopulations, and per-phasic response spike proportions. Data were combined linearly with coefficients adjusted to maximize the correlation to the behavioural data for the same training stage using Newton's nonlinear maximization method. Correlations between these scores and actual behavioural measures were computed for the entire training period and separately for

acquisition, extinction and reacquisition phases. Mathematically, the composite neural score is a linear combination of three sets of neural data: $P_i = a_1X_{1i} + a_2X_{2i} + a_3X_{3i}$, where a_1, a_2, a_3 are coefficients, X_{1i}, X_{2i}, X_{3i} correspond to values of normalized per-neuron firing rates, proportions of task-responsive neuronal subpopulations and per-phasic response spike proportions, respectively, and *i* is the index of the learning stages. The coefficients were set to the values that maximized the correlation between the composite neural scores and the behaviour accuracy at all stages. The significance of the maximum correlation was determined by computing the probability that such a correlation could arise by chance. To do so, we generated control data R_{1i}, R_{2i}, R_{3i} whose values were randomly drawn from interval (-1,1), and calculated maximum correlations to behaviour accuracy achieved with these control data. The probability distribution of the maximum correlations was constructed using 10,000 sets of control data and used to evaluate the significance of the correlation between the true composite neural score and the behaviour data.

Supplementary Discussion

We report in this study learning-related changes in the activity of neurons classified as projection neurons recorded in the dorsolateral (sensorimotor) zone of the striatum. Anatomical and behavioural evidence for functional specificity in different regions of the striatum[1-3] suggests that the observed neuronal activity patterns and their restructuring during learning could be specific to this region of the striatum, and that

projection neurons in other striatal zones could exhibit activity patterns that are different from those in the dorsolateral striatum[4].

Our findings demonstrate that these changes in the firing patterns of striatal projection neurons continued even after behavioural accuracy scores and running times approached asymptotic values during the over-training period (Fig. 2). This prolonged restructuring of neural activity despite relative behavioural stability suggests that the striatal spike activity was not a simple representation of on-going behaviour per se—at least as we were able to monitor it. The fact that the restructuring of the population firing patterns was similar for correct and incorrect trials also accords with this interpretation.

We defined non task-responsive neurons as units that did not exhibit excitatory responses that reached the level of 2 SDs above pre-trial baseline firing in per-session averaged activity in any peri-event window. Despite their classification as non task-responsive, the peak-detection algorithm that looked for peri-event increases in the firing rates of these units (with a slope threshold; Supplementary Methods) found small positive deflections in neural activity in 16.6% of these units. This result suggests the possibility that some of the non task-responsive units in fact showed or were developing task-related responses. However, the use of the peak detection algorithm, which compares the response not to pre-trial activity but to in-task activity, as the sole criterion for defining task-responsive and non task-responsive neural subpopulations yielded nearly identical patterns of neural changes (Supplementary Fig. 2). This finding supports the conclusion that firing rate changes occurred for non task-responsive as well as for task-responsive neurons in the sensorimotor striatum as behavioural learning occurred.

We performed simple sensory exams to test for the sensory responsiveness of neurons at the sites of tetrode recordings before or after selected sessions. The results of these exams did not suggest a clear relationship between the sensory receptive fields, mapped out-of-task by simple manual manipulation of joints and tactile stimulation, and the task-related activity that we observed in-task. We did, however, note unit activity related to running that had a 3-5 Hz rhythmicity (e.g., Fig. 2c top).

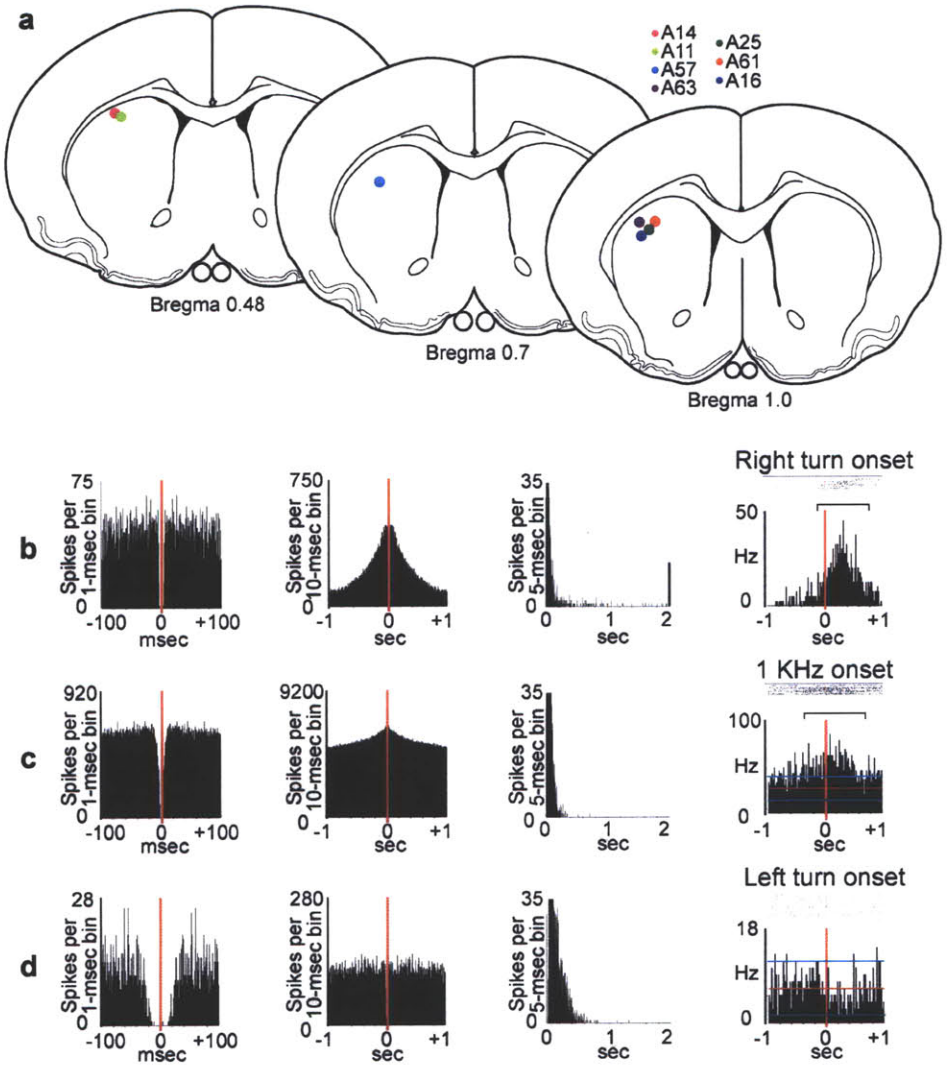
We have interpreted these findings as suggesting that the reconfiguration of activity in the sensorimotor striatum represents neural correlates of behavioural exploration-exploitation. This working hypothesis leaves open the issue of how such restructuring (e.g., temporal advances at task start and end) is generated. It is possible, for example, that the gradual lessening of the activity in the middle of the maze runs occurs independently of the temporal shifts with different versions of the maze task, and that different patterns of plasticity occur with markedly different task requirements. The fact that the multiple changes in spike patterns that we document here were expressed by neurons classified as striatal projection neurons is a critical point, because these neurons receive cortical inputs and give rise to the main outputs of the basal ganglia. It is possible that some putative projection neurons (here for convenience called projection neurons) were misclassified, as our data were acquired by extracellular ensemble recordings. From the firing properties of these neurons (Supplementary Fig. 1), however, we suspect that classification errors were minimal. The fact that the spiking patterns of striatal projection neurons were restructured with learning thus suggests that a circuit-level reconfiguration of neural activity in cortico-basal ganglia loops accompanies, and could mediate, the

initial acquisition of habitual behaviour and the adjustment of such behaviour in response to novel contexts.

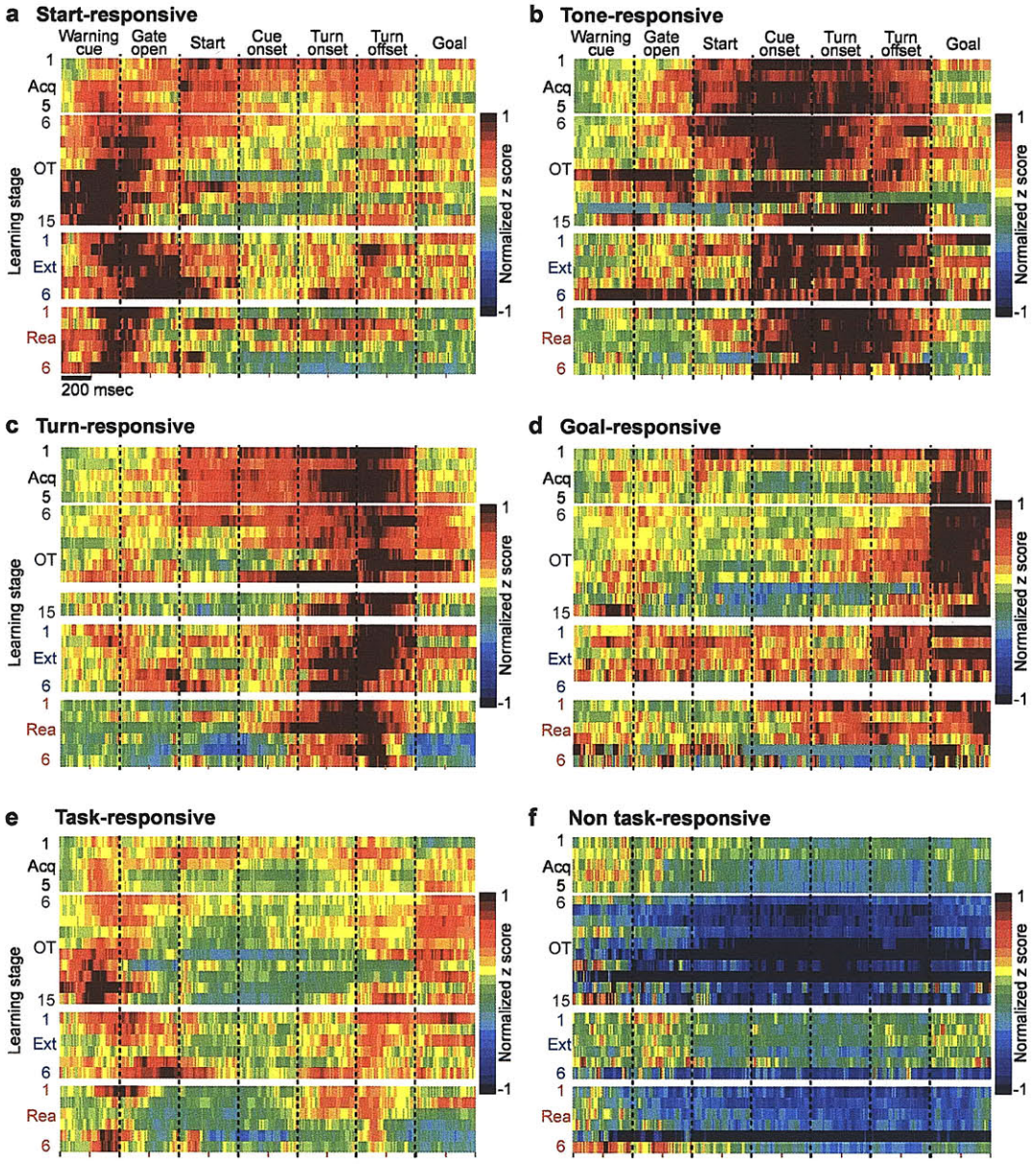
Supplemental References

1. Alexander, G. E., DeLong, M. R. & Strick, P. L. in Annu Rev Neurosci (eds. Cowan, W. M., Shooter, E. M., Stevens, C. F. & Thompson, R. F.) 357-381 (Annual Reviews, Palo Alto, 1986).
2. Devan, B. D. & White, N. M. Parallel information processing in the dorsal striatum: relation to hippocampal function. J Neurosci 19, 2789-2798 (1999).
3. Yin, H. H. & Knowlton, B. J. Contributions of striatal subregions to place and response learning. Learn Mem 11, 459-63 (2004).
4. Kubota, Y., DeCoteau, W. E., Liu, J. & Graybiel, A. M. (Program No. 765.7. 2002 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2002. Online., 2002).
5. Paxinos, G. & Watson, C. The Rat Brain in Stereotaxic Coordinates. Third Edition (Academic Press, San Diego, 1997).
6. Tepper, J. M., Koos, T. & Wilson, C. J. GABAergic microcircuits in the neostriatum. Trends Neurosci 27, 662-9 (2004).
7. Wilson, C. J. in The Synaptic Organization of the Brain (ed. Shepherd, G. M.) 361-413 (Oxford University Press, New York, 2004).

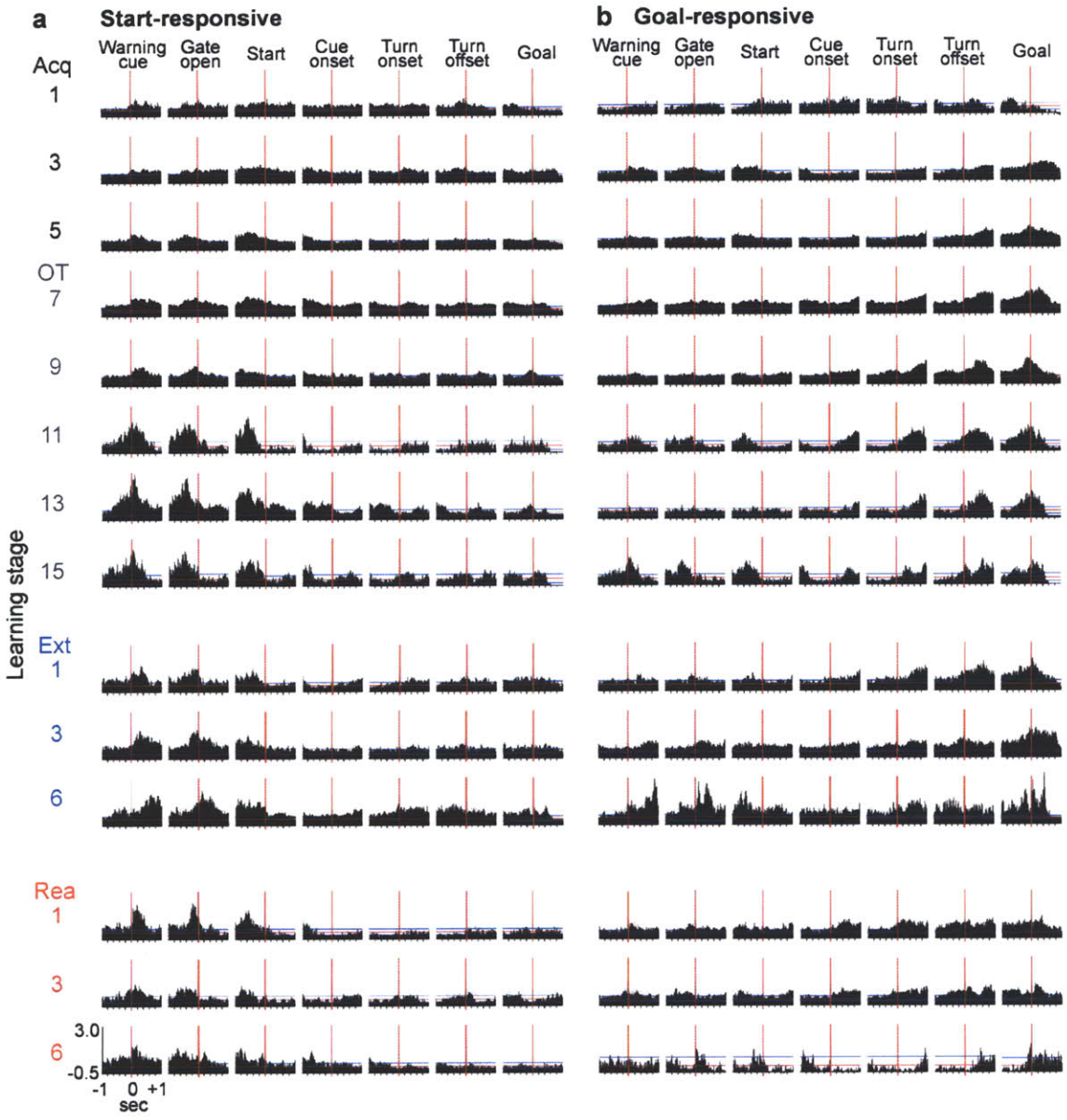
Supplementary Figures



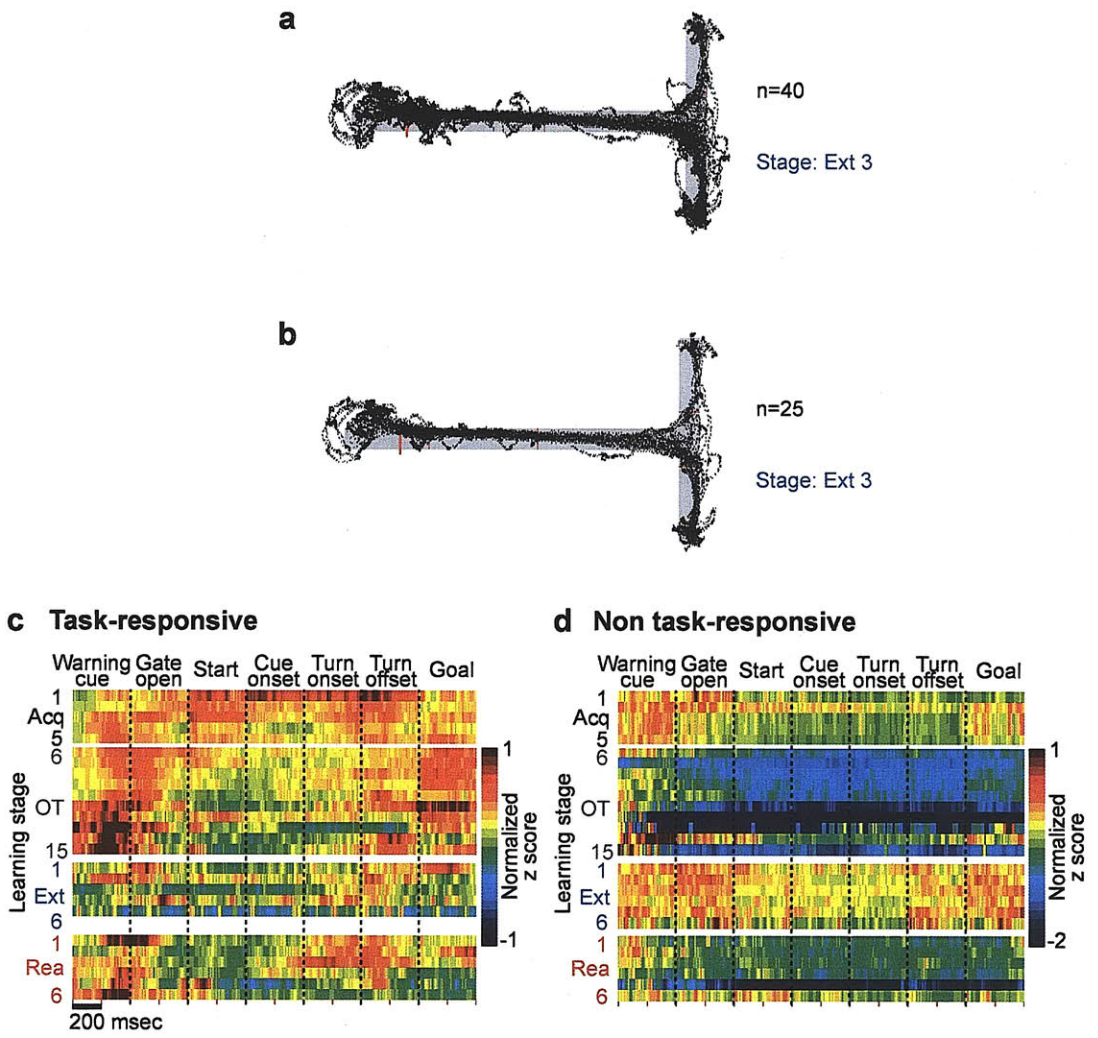
Supplementary Figure 1 Recording sites and unit classification methods. **a**, Schematic outlines[5] illustrating in color-coded dots the estimated locations of final recording sites in each of the seven rats included in this study. **b-d**, Spike patterns of typical units classified as putative projection neurons (**b**), putative fast-firing neurons (**c**), and putative tonically-firing neurons (**d**). Shown for each unit are two auto-correlograms (with ±100-msec and with ±1-sec windows, two left columns), an interspike interval (ISI) plot (middle), and a peri-event raster plot and histogram (right). Putative projection neurons were characterized by the occurrence of some long (> 2 sec) ISIs (bar at 2 sec) in addition to phasic activity with short ISIs, putative fast-firing neurons by high average firing rates and a lack of ISIs > 1 sec[6], and putative tonically-firing neurons by wide central valleys in the auto-correlograms and spiking at 4-10 Hz[7]. In the peri-event histograms shown in **c** and **d**, the mean firing rate during the pre-trial baseline period and the level 2 SDs above and below the mean are indicated, respectively, by red and blue horizontal lines (Supplementary Methods). The unit shown in **b** did not fire during the baseline period. Phasic responses (indicated by brackets in histograms) were identified by the peak detection algorithm described in Supplementary Methods for units shown in **b** (from -0.15 to +0.77 sec) and **c** (from -0.39 to +0.70), but the algorithm did not detect a phasic peak in spiking for the unit shown in **d**.



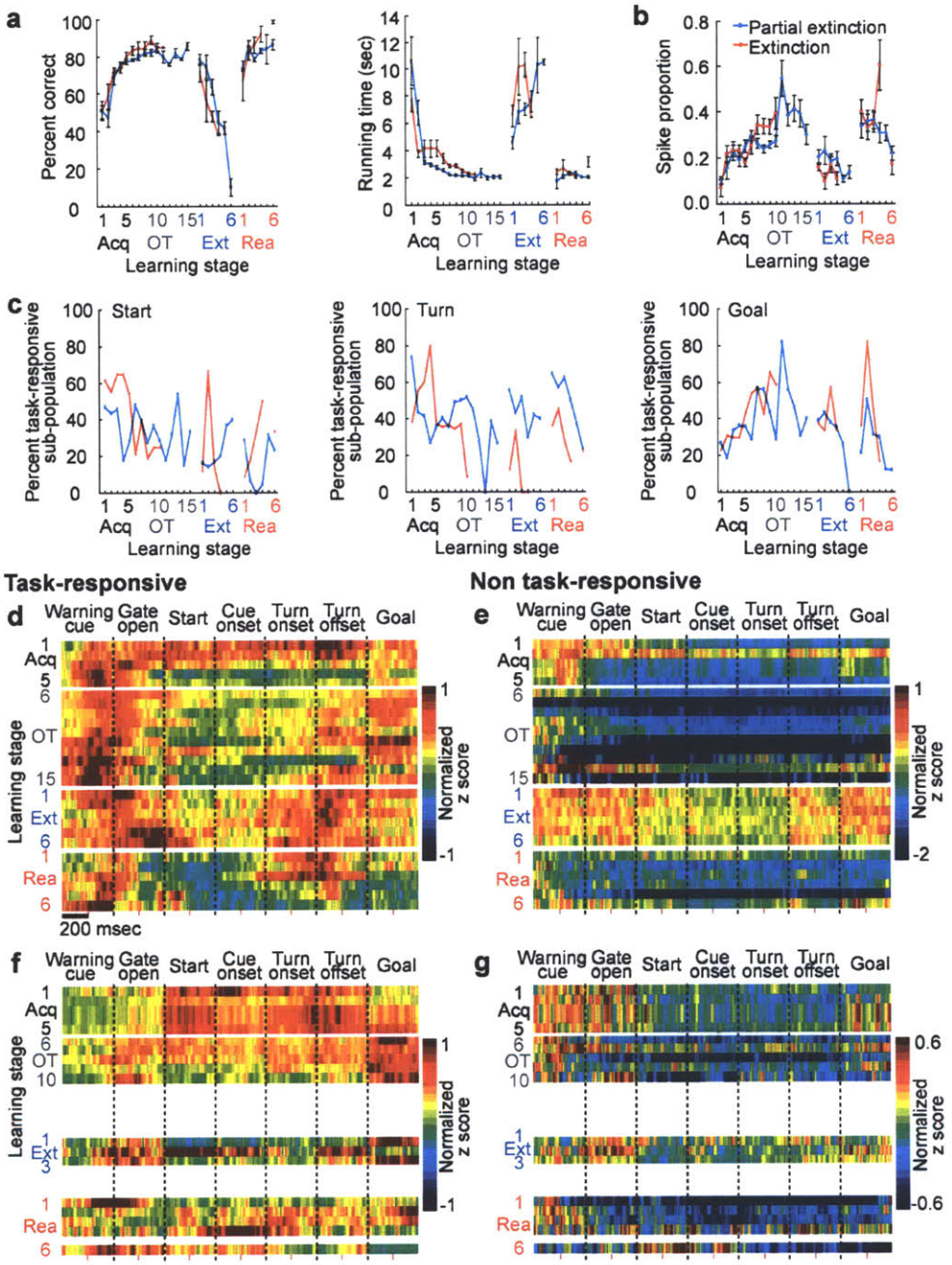
Supplementary Figure 2 Normalized averaged per-neuron spike frequency plots constructed as in Fig. 2. **a-d**, The activity patterns of subpopulations of striatal projection neurons. Each plot shows activity of neurons with start-related (**a**), tone-related (**b**), turn-related (**c**) and goal-related (**d**) responses for acquisition (Acq), over-training (OT), extinction (Ext) and reacquisition (Rea) training periods. Note that the temporal shifts in activity at task start and end exhibited by the entire populations of task-responsive projection neurons (Fig. 2) are accentuated in the neuronal subpopulations of start- and goal-responsive units. **e-f**, Population per-trial activity patterns of units classified as task-responsive and non task-responsive based on peri-event phasic responses detected by a peak detection algorithm that tested for changes in firing rates in moving time-windows irrespective of pre-trial baseline activity (see Supplementary Methods). Average activity of projection neurons for which phasic activity peaks were detected within 200-msec around task events (**e**) and those for which such phasic activity was not detected (**f**). Note similarities to the activity patterns of task-responsive and non task-responsive units shown in Fig. 2, which were classified on the basis of peri-event spike rates relative to pre-trial baseline firing.



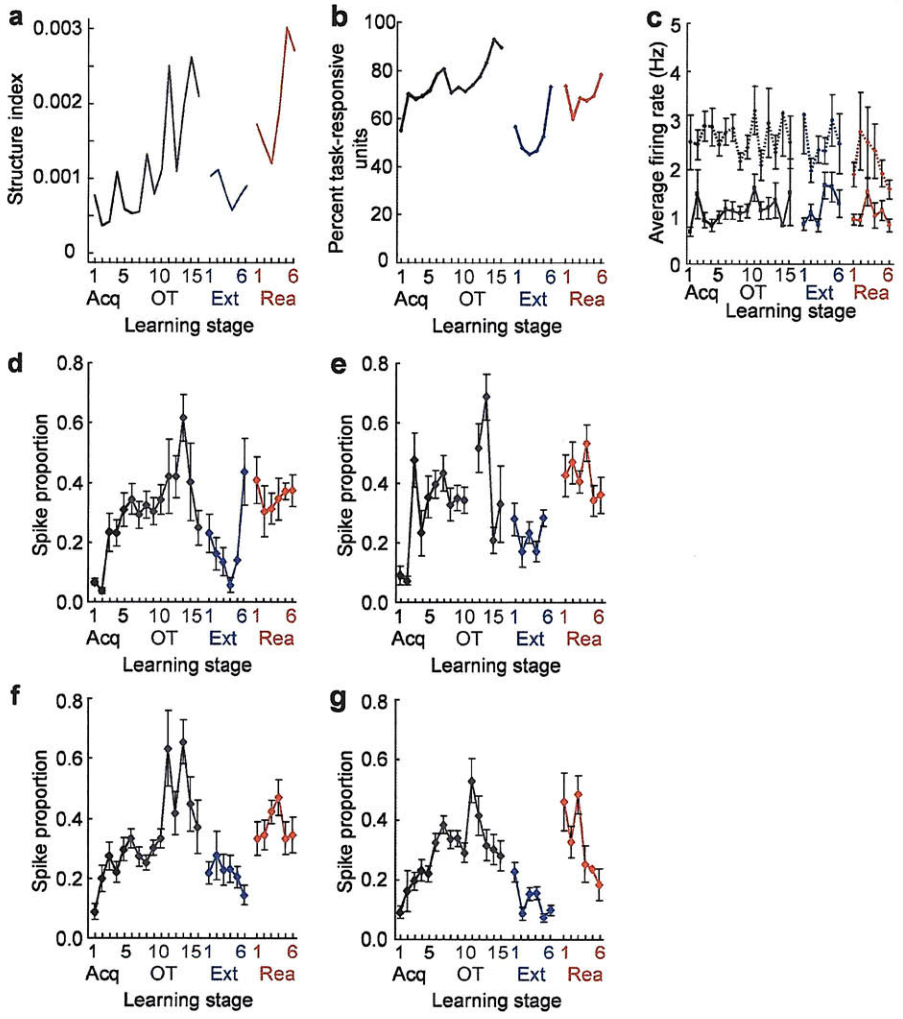
Supplementary Figure 3 Population firing patterns of striatal projection neurons. **a-b**, Activity of neurons exhibiting task-related activity near the start of the trial runs (**a**) and those with responses at goal reaching (**b**). Each plot shows activity in consecutive 20-msec bins during ±1-sec windows around task events as indicated, averaged over all units of the type. Data are shown for every second stage of acquisition, extinction and reacquisition training. Note that during acquisition (stages 1, 3, 5) and over-training (7, 9, 11, 13, 15), phasic responses gradually developed, and that these phasic activities gradually shifted earlier in time. During extinction, the temporal shifts reversed. During reacquisition training, strong activity peaks re-emerged at task start and again shifted earlier.



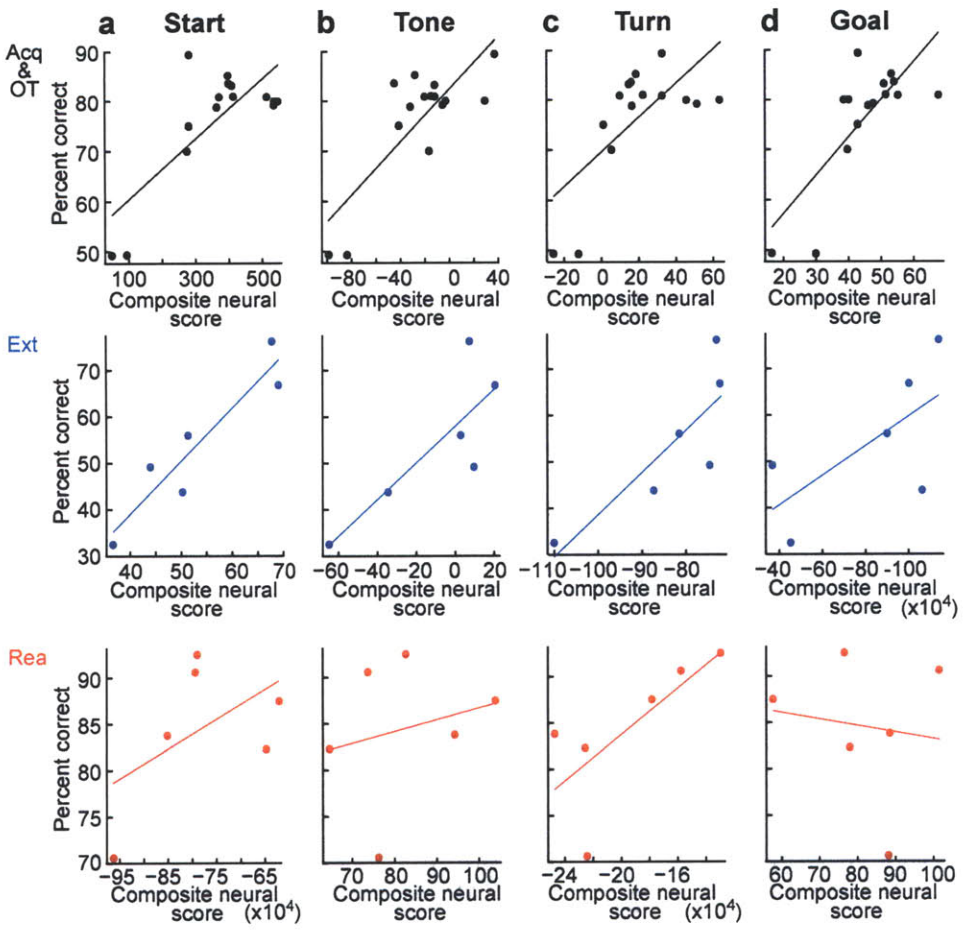
Supplementary Figure 4 Extinction-induced reversal of reconfigured striatal activity is not due to increase in running times during extinction. For these analyses, all extinction trials with running times greater than 1 SD over the mean of all trials in over-training sessions were discarded to increase behavioural stationarity of the extinction data analyzed. **a-b**, Overlaid trajectories of a rat during all trials of a single extinction session (**a**) and those during trials of the same session that were included in this analysis (**b**). **c-d**, Averaged normalized activity of task-responsive (**c**) and non task-responsive (**d**) projection neurons during all trials of acquisition, over-training and reacquisition stages and during extinction trials that met the mean +1 SD criterion. Note patterns of changes during extinction (particularly for non task-responsive units) are similar to those with all extinction trials shown in Figure 2, despite fewer numbers of trials included in this analysis.



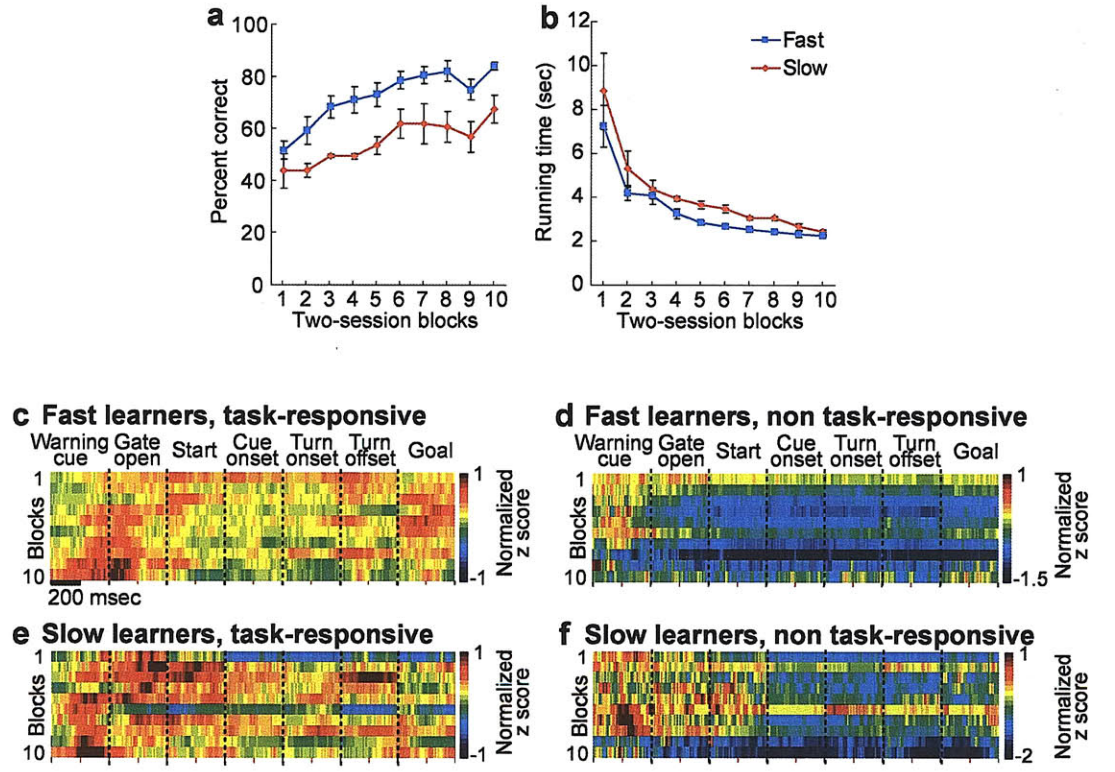
Supplementary Figure 5 Partial and full extinction training procedures yielded similar behavioural and neural changes. **a**, Percentage of correct responses (left) and running times (right) for rats trained on the partial extinction procedure (n = 4, blue) and those trained on the full extinction procedure (n = 3, red). Rats given partial extinction training received reward in 1-3 trials per session, whereas no reward was given to rats trained on the full extinction procedure. Error bars indicate standard errors of the mean. **b**, Average proportions of spikes that occurred within the periods of phasic responses relative to the spikes that occurred during the entire trial time, averaged for each session. **c**, Percentages of projection neurons that exhibited responses at task start (left), turning (middle) and goal reaching (right). **d-g**, Activity of projection neurons with and without phasic task-related responses, recorded in rats receiving the partial extinction procedure (**d** and **e**, respectively) and in rats receiving the full extinction procedure (**f** and **g**, respectively). Patterns in **f** and **g** were not as clean as in **d** and **e**, given the lower numbers of stages attained.



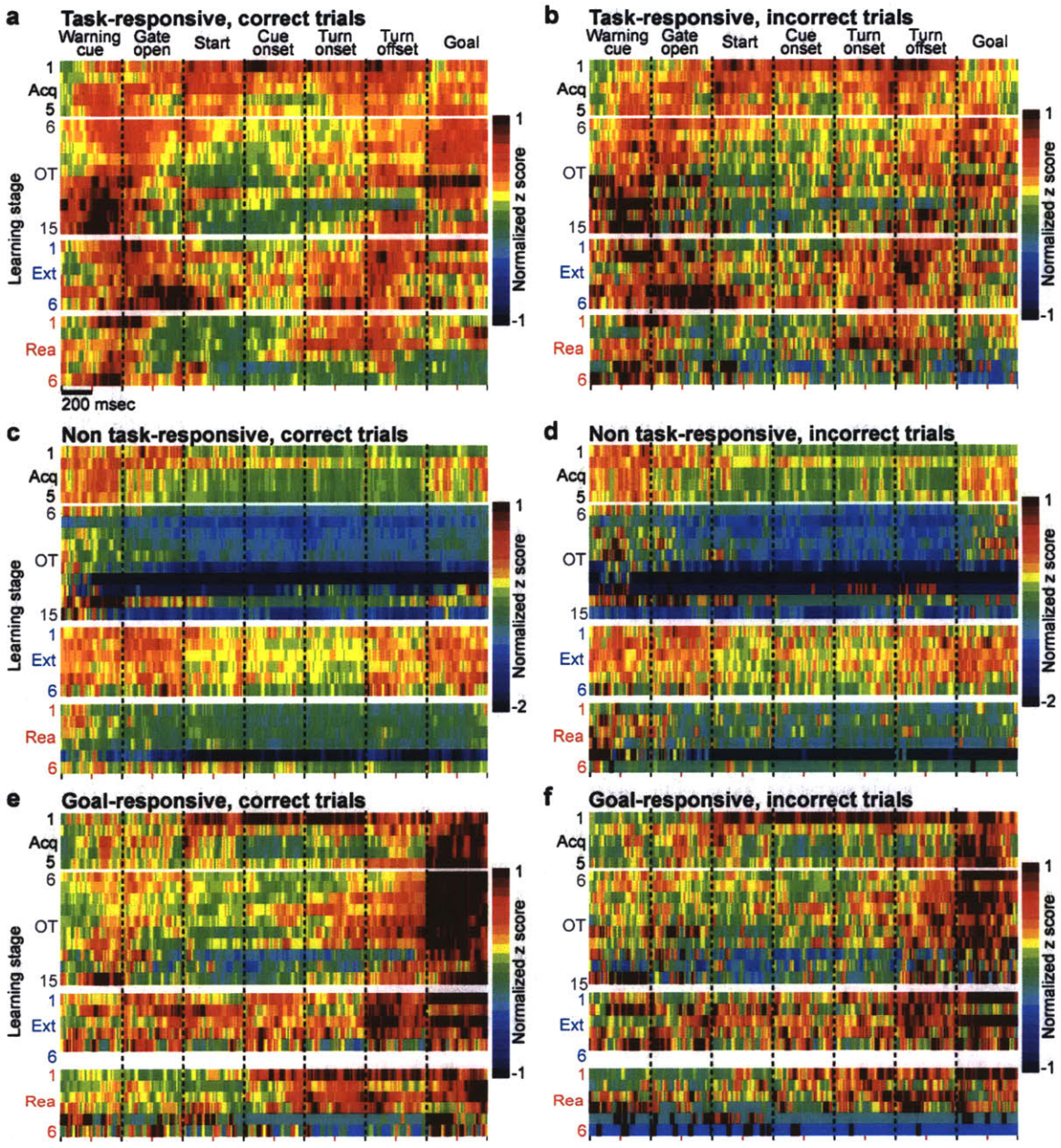
Supplementary Figure 6 Multiple changes in striatal projection neuron activity during successive acquisition, extinction and reacquisition training. **a**, Structure in the per-trial activity of task-responsive units (Fig. 2a) calculated for each training stage (indexed by variance in spiking across the entire trial runs), shown for each learning stage. **b**, Changes in the proportions of accepted projection neurons exhibiting spike firing that met the mean + 2 SD criterion for task-related activity across stages of training. **c**, Average firing rates of units with task-related activity (shown by dotted lines) and average firing rates of units that did not have peri-event activity above the threshold during trials (shown by solid lines), plotted for the successive training stages. Error bars represent standard errors of the mean. **d-g**, Average proportions of the spikes per neuron that occurred during phasic responses to warning cue (**d**), to tone (**e**), at turn (**f**) and at goal-reaching (**g**), relative to the spikes that occurred during the entire trial time. Note that the per-neuron spiking in relation to each of these task events became concentrated in phasic responses during acquisition, even for task events in the middle of trail time, during which population spiking activity decreased during acquisition and over-training (e.g., turn). The spiking then became less concentrated in phasic responses during extinction, and then again became concentrated in phasic responses during reacquisition.



Supplementary Figure 7 Prediction of behavioural accuracy by a composite neural score combining normalized per-neuron firing rate, proportion of different task-responsive subpopulations, and per-phasic response spike proportion. **a-d**, The neural parameters were computed for ±200-msec windows around warning cue (**a**), tone onset (**b**), turn onset (**c**) and goal-reaching (**d**) for all projection neurons (Supplementary Methods). Correlations between the two measures are illustrated separately for acquisition and over-training (black, top), extinction (blue, middle) and reacquisition (red, bottom). Significant correlations were found between the neural composite scores and actual behavioural accuracy across all training stages for warning cue ($R = 0.69$, $P < 0.001$), tone ($R = 0.77$, $P = 0.009$), turning ($R = 0.69$, $P = 0.019$) and goal-reaching ($R = 0.63$, $P = 0.033$). Correlations were higher for initial acquisition and over-training ($R = 0.69 - 0.80$, $P = 0.04 - 0.07$) and extinction ($R = 0.66 - 0.93$, $P = 0.21 - 0.46$) phases of training than for reacquisition training ($R = -0.13 - 0.73$, $P = 0.30 - 0.88$).



Supplementary Figure 8 Late development of restructuring of striatal spike patterns in rats that learned the task at slow rates. **a-b**, Percentage of correct responses (**a**) and running times (**b**) during the first 20 acquisition training sessions, shown in 10 blocks of two consecutive trials, for rats that acquired the task at fast rates (blue, average number of sessions to reach the first of 10 out of 11 consecutive sessions above criterion = 10.4, n = 5) and those that acquired the task at slow rates (red, average number of sessions = 43.5, n = 2). **c-f**, Averaged normalized activity of task-responsive projection neurons recorded in fast learners (**c**) and in slow learners (**e**), and of non task-responsive neurons recorded in fast (**d**) and slow (**f**) learners. Note slow development of the learning-related patterns shown in Fig. 2 in rats that acquired the task slowly (**e** and **f**).



Supplementary Figure 9 Normalized averaged per-neuron spike frequency plots for trials with correct and incorrect behavioural responses. **a-b**, Activity patterns of all task-responsive projection neurons during trials in which performance was correct (**a**) or was incorrect (**b**). **c-d**, Activity of neurons whose firing rates did not reach the baseline mean + 2 SD criterion for acceptance as task-responsive neurons, plotted for correct (**c**) and incorrect (**d**) trials. **e-f**, Activity patterns of projection neurons that responded at goal reaching shown for trials with correct behavioural performance (**e**) and those with incorrect performance (**f**). Note the similarity between correct-trial and incorrect-trial firing patterns for all these subpopulations of projection neurons.

Supplementary Table 1 **Summary of training schedules and recording yield for the 7 rats included in the study**

	Subjects						
	A11	A14	A16	A25	A57	A61	A63
Number of Sessions	63	63	38	41	30	32	24
Acquisition	21	12	12	6	6	5	3
Over-training	38	35	16	18	10	16	12
Extinction	2	11	5	7	4	4	5
Reacquisition	2	5	5	10	10	7	4
Total Number of Accepted Units	461	635	714	579	115	244	401
Accepted Units Per Session	7.3	10.1	18.8	14.1	3.8	7.6	16.7
Average Total Tetrode Movement (mm)	0.69	0.85	0.34	0.28	-0.13	0.75	0.07

Supplementary Table 2 **Training sessions included in each learning stage**

		Subjects						
Stage #	Behavioural Criterion	A11	A14	A16	A25	A57	A61	A63
Acq 1	1st session	1	1	1	1	2	3	2
Acq 2	2nd session	2	2	2	2	3	4	3
Acq 3	> 60%	11	10	11	6	4	5	3
Acq 4	> 70%	21	12	11	6	4	5	3
Acq 5	>72.5% 1+2	21+22	12+13	12+15	6+12	6+7	5+8	3+5
OT 6	>72.5% 3+4	23+26	15+18	18+20	15+16	8+9	9+12	6+8
OT 7	>72.5% 5+6	29+30	19+20	21+22	17+18	10+11	13+14	9+10
OT 8	>72.5% 7+8	32+34	25+26	23+24	19+20	12+14	15+16	11+12
OT 9	>72.5% 9+10	35+38	28+29	25+26	21+22	15+16	17+18	13+14
OT 10	>72.5% 11+12	41+42	30+32	27+28	23+24		19+20	15
OT 11	>72.5% 13+14	45+46	33+34				21	
OT 12	>72.5% 15+16	47+49	38+39					
OT 13	>72.5% 17+18	50+51	40+41					
OT 14	>72.5% 19+20	53+54	42+43					
OT 15	>72.5% 21+22	55+56	44+45					
Ext 1	1st session	60	48	29	25	17	22	16
Ext 2	2nd session	61	49	30	26	18	23	17
Ext 3	3rd & 4th sessions		50+51	31+32	27+28	19+20	24+25	18+19
Ext 4	5th & 6th sessions		52+53	33	29+30			20
Ext 5	7th & 8th sessions		54+55		31			
Ext 6	9th & 10th sessions		56+57					
Rea 1	1st session	62	60	34	32	21	26	21
Rea 2	2nd session	63	61	35	33	22	27	22
Rea 3	3rd & 4th sessions		62+63	36+37	34+35	23+24	28+29	23+25
Rea 4	5th & 6th sessions		64	38	36+37	25+26	30+31	
Rea 5	7th & 8th sessions				38+39	27+28	32	
Rea 6	9th & 10th sessions				40+41	29+30		

Chapter Three

Firing Patterns of Striatal Neurons Reflect Chunking of a Motor Program

Abstract

In order to more fully understand the role the sensorimotor (dorsolateral) striatum plays in motor sequence learning and habit formation, we trained rats on a discriminative T-maze task in which tones played at the beginning of the task were predictive of which end arm was baited with chocolate. Animals were therefore given all the information needed to pre-plan the motor sequence through the maze, which contrasts with our previous version of the discriminative T-maze which provided cues at mid-run. We recorded throughout the entire learning process with tetrodes in order to capture the learning related changes that occurred in firing patterns in the sensorimotor striatum. We found a strong task bracketing pattern of activity develop wherein the average firing rate of task responsive cells in the striatum increased at the beginning and end of the task and decreased during the middle of the task. Furthermore, we determined that what is actually bracketed is the motor sequence, not the entire trial, as the increase in firing at the beginning of the task does not occur during the discriminative tone (the first part of every trial). The pattern became more consistent from trial to trial within single cells throughout learning and showed differences during certain types of trials (e.g. correct and incorrect trials). When contrasted with previous data in a task using mid-maze cues, this pattern was extremely strong. We conclude that giving the animal all of the necessary information at the beginning of the task allows the animal to preplan the entire motor sequence or habit and may increase the strength of chunking that occurs in the striatum. This preplanning may be what causes the strengthening of the observed task bracketing pattern.

Introduction

The dorsolateral striatum, part of the major sensorimotor input structure of the basal ganglia, is thought to be involved in habit formation and sequence learning (Packard and McGaugh 1996; Barnes, Kubota et al. 2005; Bailey and Mair 2006; Yin and Knowlton 2006; Atallah, Lopez-Paniagua et al. 2007; Belin, Jonkman et al. 2009). Damage to the dorsolateral striatum causes deficiencies in habit and sequence task learning (Packard and McGaugh 1996; Bailey and Mair 2006; Yin and Knowlton 2006; Atallah, Lopez-Paniagua et al. 2007; Belin, Jonkman et al. 2009). There is some evidence to suggest that the dorsolateral striatum is not needed to learn sequences, but it is necessary to perform learned sequences. (Atallah, Lopez-Paniagua et al. 2007). Firing patterns in the dorsolateral striatum have been shown to change systematically as animals are trained in a variety of sequence learning tasks to encode action automation and habit formation (Chang and Gold 2003; Costa, Cohen et al. 2004; Schmitzer-Torbert and Redish 2004; Yin, Knowlton et al. 2004; Barnes, Kubota et al. 2005).

Not unlike memorizing phone numbers, sequences are often learned in behavioral units called chunks (Miller 1956; Graybiel 1998; Sakai, Kitaguchi et al. 2003). For example, large sequences of learned motor taps are performed in subsequences, or chunks, of taps separated by time gaps (Rosenbaum, Kenny et al. 1983). By allowing an entire complex sequence to be initiated by the first item in that sequence, performing complex behaviors becomes more automatic. It has been proposed that the dorsolateral striatum and related structures are involved in the chunking of sequences of smaller motor movements into larger motor chunks (Graybiel 1998; Sakai, Kitaguchi et al. 2003).

Previously, we found that a pattern develops in the spiking rates of cells recorded in the dorsolateral striatum that was consistent with this chunking hypothesis (Barnes, Kubota et al. 2005). Rats first heard a click that was followed by a gate opening. Next, rats ran down the long arm of the T-maze. Mid run, one of two tone cues turned on. Each tone was associated with one of the two goal arms that was baited. Rats learned to associate the tone heard with the goal arm that was baited. In this task we found that early in training cells fired, on average, throughout the trial. As training progressed, more firing occurred during the beginning and end of the trial, as though to chunk the action sequence (Jog, Kubota et al. 1999; Barnes, Kubota et al. 2005).

In the current experiment, by moving the discriminative tone to the beginning of the task and thereby supplying the animal with all information needed to plan the entire motor sequence at the start of the trial, we manipulated the task to evaluate motor chunking features. If chunking is indeed responsible for the increase in firing at the beginning and end of the task, we expected to see a stronger chunking pattern develop in the dorsolateral striatum as learning progressed on this task.

Results

At the beginning of each trial, a tone –1 kHz or 8 kHz- was played that signaled which arm of the T-maze would contain a reward (**Figure 1a**). Rats learned in a series of 40 trial sessions to traverse the maze and select the rewarded arm. The averaged staged percentage of correct trials per session across the population of rats began at chance performance (50%) and increased to above the predetermined ($p<.01$) 72.5% criterion for

learning as training progressed (**Figure 1b** $p < .0001$). The average amount of days it took for animals to first reach our criterion was 15 days. However the percentage of correct trials, on a session-to-session basis for an individual rat was variable. Therefore, we also estimated the onset of learning using the dynamic state-space model paradigm developed by Smith et al. (Smith, Frank et al. 2004), (**Supplemental Figure 1**). This model characterizes learning as the probability an animal will maintain a higher than chance performance for the duration of the experiment. Estimated this way it took an average of 8.5 sessions (± .56 SEM) for the animals to learn this task (excluding one animal that, according to this model, failed to learn). Running times also significantly decreased with training (**Figure 1c** $p< .001$).

Population firing rates

The average spike rate of task-responsive medium spiny neurons recorded from the dorsolateral striatum across training stages restructured dramatically as animals learned this task (**Figure 2a**). Average per-neuron spike rates increased at the beginning (after gate opening) and end (as animals approached reward) of trials while, concomitantly, the average spike rate decreased in the middle portion of trials (**Figure 2b**). This restructuring of neuronal firing patterns was progressive and happened gradually throughout the course of learning (**Figure 2c**). In contrast, non-task responsive neurons, as an ensemble, during the trial, maintained their firing rate as training progressed. Unlike task responsive cells, the firing rate for non-task responsive cells during the baseline period increased throughout the trial (**Figure 4**).

This progressive accentuation of firing of task-responsive cells at the beginning and end points of the maze is not due to how we averaged across cells. It can be seen in

the averaged firing rates normalized with respect to different baselines and calculated in different ways (**Supplemental Figure 2**). Indeed it is so strong that it can also be seen in the firing patterns of single cells viewed as an ensemble (**Figure 3).** It is important to note that although the average firing rate of task responsive cells grew to accentuate the beginning and end of each trial, all cells did not follow this pattern. Cells responded during virtually all times of the trial analyzed (**Figure 3**).

Source of population increases and decreases

In order to determine if this progressive annunciation of the beginning and end of each trial run was due to 1) an increase or decrease in the amount of cells firing at these times or 2) similar proportions of cells modifying their firing rate, we examined 400-millisecond time windows during the beginning, middle, and end of the task and determined whether the percentage of cells hitting their maximum firing rate or 25% of their maximum firing rate increased or decreased as learning progressed. We also determined what this level of firing was and whether it changed as training progressed.

The increase in average population firing rate at gate can be attributed to a relatively consistent percentage of single cells that increased their firing rate. The percentage of cells reaching their maximum rate during the 400 milliseconds after gate opening averaged 12% across learning and did not increase as training progressed. There was a nominal increase in the size of this maximum (ANOVA stage 1 and stage 9 $p < .1$ **Figure 5**). The percentage of cells reaching 25% of their maximum firing rate during this time period averaged 51% and did not change as training progressed. However, the firing rate for those cells that reached at least 25% of maximum firing rate increased from

3.7 Hz to 8.4 Hz (ANOVA stage 1 and 9 $p < .01$ (**Figure 5**) demonstrating that this subset of cells increased their responses to gate opening as training progressed.

In contrast to changes in firing during the gate event, the decrease in the average firing rate during the middle of the trial as training progressed can be attributed to a fewer percentage of cells firing during this time period. There were very few cells , and average of 2%, that reached their maximum firing rate during the middle of the run (400 milliseconds after the mid-run photobeam). The decrease in the magnitude of the 25% maximum firing rate was not significant, however, the percentage of cells reaching at least 25% of their maximum during this time decreased as training progressed from 42% to 4% (ANOVA $p < .05$ **Figure 5**).

At goal arrival, both an increase in the percentage of cells responding to goal and an increase in firing rate of cells that responded to goal occurred as training progressed. The percentage of cells reaching their maximum firing rate during the end of the trial runs (the 400 milliseconds after reaching the goal) increased as training progressed from 15% to 44% (chi square test $p < .01$, **Figure 5**). The amplitude of the maximum also increased with training (ANOVA $p<.05$) from an average of 9.6 hertz to an average of 11.8 hertz. The amplitude of 25% of the maximum also increased from 5.4 Hz to 7.9 Hz (ANOVA $p<.05$).

Consistency of neural firing patterns

We next asked whether firing patterns of single cells across trials were becoming more consistent as training progressed. In order to address this question we correlated the firing pattern of a cell during each trial with the firing pattern of every other trial. We

found that the correlations of firing patterns of trials increased from an average correlation of .10 on stage one (pre-learning) to .21 on stage nine (well-learned). This demonstrates that cells became more consistent as training progressed (ANOVA P< .0001 **Figure 6a).** Trial by trial variability in firing patterns around gate in particular became more consistent as training progressed (**Figure 6b**). This was not true for the pre-trial baseline period (ANOVA p >.05), where the correlations were unchanged with learning.

We also analyzed trial by trial consistency by looking at the consistency of firing of single cells across the entire ensemble. Forty-five cells in each stage were selected (which was the lowest number of cells in any stage). For each cell, we took 800 ms around each task event (including a pre tone baseline) and then determined where the maximum firing rate occurred in the trial. We then ranked the cells recorded in each stage according to when this maximum occurred in relation to the other 44 cells from this stage. We did this for 1000 bootstraps with this data and the shuffled data and then took the standard deviation of rank. The lower the standard deviation in rank, the more consistently the cell fired across trials. Shuffled data received a standard deviation of rank lower than 8 less than 1% percent of the time. On the first day of training, 44% of cells analyzed showed a standard deviation of rank lower than 8. On the last stage of training 78% of cells showed this lower standard deviation of rank (Chi Square p<.001 **Figure 6c** and **Figure 6d**). Therefore, the amount of cells showing more consistency from trial to trial increased compared to chance across learning.

We also examined the consistency of the changes in neuronal firing pattern from animal to animal. Aspects of the pattern could be seen in the majority of animals.

However, the strength and type of learning related changes that occurred in the spiking patterns of dorsolateral striatal neurons varied from animal to animal (**Supplemental Figure 3**). Supplemental **Figures 3b, c and d** show the learning curve, running time and average neuronal firing pattern of task responsive cells for an animal with a strong accentuation of the beginning and end that learned the task.

There were several different types of trials that occurred during each session. In some trials animals chose correctly and therefore received a chocolate reward. In others, incorrect choices were made and no reward was given. We next examined whether the neuronal firing patterns in the dorsolateral striatum of well trained animals (stage five and higher) changed in correct trials compared to incorrect trials. We found that the average normalized firing rates in correct and incorrect trials were remarkably similar until rats approached goal (**Figure 7a**). As rats approached the goal, more firing was observed in correct trials. We also examined whether firing patterns were different for trials immediately following a correct trial compared to trials following an incorrect trial in well trained animals (**Figure 7b**). We found no difference between these types of trials.

After rats reached the behavioral performance criterion of 72.5%, they did not always maintain above chance performance for a given session. We could therefore analyze neuronal firing patterns for sessions after rats had learned our task in which rats performed poorly (at or below 60%) and compare these sessions to sessions in which the rats performed well (at or above 70%). Interestingly, we found that there was a decrease in the average firing rate on good performance days in the 400 milliseconds before gate opening and again at goal.

Relationship of neural firing to speed

Next, we examined the relationship between the firing rates in dorsolateral striatal neurons and the speed of rats as they ran the maze in order to determine if, in part, the learning related neuronal changes seen in the dorsolateral striatum could be attributed to decreases in run time as training progressed. The average speed across the trial throughout learning is plotted in **figure 8a**. We correlated the average speed and average firing rate for all 800 millisecond peri-event time windows and found that the correlation was inconsistent starting at an r = 0.3 on stage 1 and falling to a low of r = -.54 on stage 7 (**Figure 8b**). However, it is important to note that in this task, location and speed are highly correlated, and are entangled with a range of salient task events.

In a cell by cell analysis, a large number of cells showed a significant, but modest, correlation with speed compared to shuffled data. Three hundred and four cells out of three hundred and thirty six cells showed a significant correlation. Speed accounted for as much as 10% of the variability in the activity of the most highly correlated cells. However, on average across all cells in this data set, speed accounted for less than 1% of the variability found in firing rates of single cells. Five percent of cells had a correlation high enough to account for 5% of the variability in the sample. Excluding these cells from analysis did not destroy the pattern of training induced changes seen in the striatal firing patterns.

To further determine if changes in speed throughout learning underlie the changes seen in firing rates, we looked at the changes in firing rates when speed was held constant. We determined the average speed and its standard deviations (SDs) for sessions

included in stages 5 through 8 (those stages where the average speed and average firing rate were consistently negatively correlated) and we excluded trials in stages 1 though 9 that were lower than one standard deviation below this mean (**Figure 8d).** We matched the number of trials in each stage by randomly selecting 246 trials for each session (the lowest amount of trials for any session that fit the criteria). Therefore, in this data set, the speed is similar for all of the trials plotted. Note that the pattern of ensemble firing still changed with learning. This outcome demonstrates that the changes in the firing patterns seen here are dissociable from changes in the speed of the animals; when the speed was held constant, the pattern still emerged over learning.

Lastly a multilinear regression was performed for each stage of learning with speed, acceleration and position as variables. **Supplemental Figure 4** shows the variation in the adjusted R squared values during learning and show that they are similar for position and speed.

Comparison of firing patterns in different tasks

We directly compared firing patterns recorded in this version of the T-maze task to a prior version. Restructured spike activity after learning exhibited significant differences compared to the restructured spike activity detected in a previous study, in which the tone cue was instead presented in the middle of the maze and remained on for the duration of the trial (**Figure 9**) (Jog, Kubota et al. 1999; Barnes, Kubota et al. 2005). This contrasted with our current task in which the tone was presented transiently (for 500 ms) prior to opening of the maze gate, requiring the rat to maintain tone identity in memory for correct arm entry and trial performance. As tone location/duration was the

only distinguishing variable between these task versions, direct comparison of spike plasticity between these groups of animals allowed us to test the hypothesis that giving all the necessary information to the animal at the beginning of each trial would allow the striatum to more strongly chunk the entire motor sequence and therefore would strengthen the beginning and end pattern. This is indeed what we found in a direct comparison of learning-related spiking patterns between animals run on these two task versions. **Figure 9d** shows raw average activity during learning stage 5 through 9 (stages where the animal performed at 72.5 % correct or higher). Learning the present task strengthened the pattern of firing that developed to bracket task performance compared to the pattern found in the previous study, primarily by decreasing the average firing rate in the middle of the trial. Interestingly, the increase in activity which developed to mark the beginning of the trials occurred at different times in the two versions of the task. The increase in activity occurred in response to the warning click (which was directly followed by the gate) in the version of the task in which the tone occurred in the middle of the trial and occurred in response to the gate in this task (where the tone occurs at the beginning of each trial). Lastly, in the present task we detected an increase in activity (compared to the activity preceding when the animal turns) at the point when the animal turns, though this activity did not change as training progressed.

In this task, the tone played for 500 ms at the beginning of each trial and then was turned off. Therefore, in order to perform this task correctly the animal was required to hold in short term memory either the identity of the tone that played or the planned turn direction. We therefore looked in the striatum in order to see if short term memory was present during this memory interval. We found no increase in the amount of cells

increasing their firing rate during the interval where the animal must be remembering the information as training progressed.

Discussion

We found a strong beginning and end pattern of the average firing rate of cells in the dorsolateral striatum that developed when animals learned this discriminative T-maze task. Early in training, on average, cells fired throughout the task. During training more firing occurred at the beginning and end of the trial runs. Changes in the percentage of cells (end) as well as the amount of spikes per cell (beginning and end) led to the accentuation of firing at the beginning and end of trials.

This pattern was markedly more pronounced when compared with our previous experiment wherein the tone sounded during the middle of the task. The primary difference between the two tasks is that in this task the tone occurs at the beginning of every trial instead of during the middle of each trial, providing the animal all the necessary information to preplan the motor sequence for a given trial. (Barnes, Kubota et al. 2005). This could allow the chunking functions of the corticostriatal circuit to be more strongly engaged. This task is also more difficult. Therefore, either the difficulty of the task or the difference in the amount of reward obtained could, in part, explain the differences in the neuronal firing patterns in the two tasks. The strengthening of the beginning and end pattern may also be an indirect result of the short term memory component to this task, although we failed to find direct evidence of short term memory in cells of the striatum. It may also be that this task causes stronger inhibition of the striatum by another structure.

Interestingly, this pattern seemed to bracket the motor portion of our task. The increase in average firing rate occurred after gate opening (the signal to animals that they could start to run towards the goal). The increase did not immediately follow tone onset, which implies that the beginning activity does not mark the earliest predictor of reward but rather seems to mark the earliest predictor of the movement period (Schultz, Dayan et al. 1997). This is unsurprising given the motor nature of much of the afferent input into the striatum (McGeorge and Faull 1989). Since runtimes decrease with training on motor tasks, one may ask if the systematic changes seen in the striatum over learning are caused by systematic changes in speed. Indeed, a high correlation between firing rates in some cells in the medial striatum and running speed of the animal has been found in previous experiments (Yeshenko, Guazzelli et al. 2004; Eschenko and Mizumori 2007). In our task, however, the systematic changes seen across learning in the average firing rate of dorsolateral striatal cells does not appear to be caused by changes in speed or acceleration alone (see **Figure 8**).

Animals had more difficulty learning this task compared to the previous task. Once our behavioral criterion was met rats had several low performing days. Yet the striatal firing patterns seen to develop with learning was stronger. Therefore, it does not appear that the ease of the task leads to a stronger pattern, rather the complexity of the task seemed to tap the striatal circuit more strongly, leading to a stronger chunking pattern.

Increases in the average firing rate occurred at the beginning and end of the task. There were, however, cells in all stages of training that responded to virtually all aspects of our task. Notably, we found few single cell responses to our mid run photobeam, a

time in the task where no events (e.g. tone on, turn on, goal) occurred (see **Figure 3**). This is consistent with the idea that cells in the striatum do not encode place but rather the order of events in a task (Schmitzer-Torbert and Redish 2008). However, similar the hippocampus, only a subset of cells are involved in the coding of any particular task (Wilson and McNaughton 1993).

We found that the baseline of non task responsive cells increased as training progressed. This is in contrast to the baseline of task responsive cells which was unchanged by learning. This may be because non task responsive cells, as training progresses, learn to encode the pre task trial time or anticipate the beginning of the task.

The gross pattern of firing in the striatum did not appear to change during correct and incorrect trials except at goal where it increased in correct trials. This is consistent with the actor/critic model of learning (Barto 1995). Interestingly, on days when animals preformed well compared to days when they did not, there was a slight inhibition just before the increase in average firing rate at the beginning of the trial. This may be due to attention effects.

We found that single cells became more consistent from trial to trial as training progressed. This is consistent with an explore-exploit model of learning in the basal ganglia (Barnes, Kubota et al. 2005) as well as with studies from the bird song field that suggest that the basal ganglia may be involved in causing perturbations in song (Kao, Doupe et al. 2005; Olveczky, Andalman et al. 2005).

The fact that the averaged, normalized firing rates were so similar until goal in correct and incorrect trials is also consistent with the actor/critic model of the basal

ganglia. The firing patterns in the dorsolateral striatum did not seem to contain more information than the animal demonstrated.

We failed to find any clear evidence of short term memory in the striatum. There were more cells that decreased their firing rate during the memory period in this task compared to a previous version that had no obvious short term memory component. However, this can be explained in many ways. The most likely explanation is that many of the cells fitting our criteria for exhibiting short term memory firing patterns were non task cells that showed a decrease during the memory period. As seen in **Figure 4**, we found that the baseline of non task responsive cells increased as training progressed. Since our criteria for showing short term memory was a decrease or increase compared to baseline, these cells may not have been decreasing their firing rate during the memory period as much as increasing their firing rate during the baseline period.

Another possible explanation is that in the non-short term memory version, the tone sounded mid run. Cells in that task may simply be more engaged in anticipating and attending to the tone. We also failed to find a difference in trials that followed a correct trial compared to trials that followed an incorrect trial.

Methods

Subjects, Habituation, and Surgical Procedures. All procedures met the approval of the Massachusetts Institute of Technology Committee on Animal Care and were in accordance with the National Research Council's Guide for the Care and Use of Laboratory Animals. Male Sprague Dawley rats (n=7; 300-350 g) were first handled for 3-5 days and then acclimated to the T-maze. Acclimation entailed sparsely scattering

chocolate uniformly around the maze and allowing the rats to freely explore the environment. Once acclimated, rats were given up to 10 'pretrials' prior to surgery in which the rat was placed in the start location, the gate was lowered, and the rat was allowed to traverse the maze. No auditory tones were played. During these pretrials the chocolate was initially scattered throughout the maze, then placed only in the food wells.

For surgery, rats were pre-treated with atropine (0.06 mg/kg) and anesthetized with ketamine (75-100 mg/kg) and xylazine (10-20 mg/kg). Small burr holes were made in the skull above unilateral dorsal-lateral striatum (AP = +0.5 mm; ML = + 3.6 mm), and the underlying dura was carefully removed. A head stage containing 7 independently moveable tetrodes (200-250 KΩ) made of twisted 10-μm Ni/Cr wire (Kanthal Palm Coast, Palm Coast, FL) was lowered into the holes to the level of dura, and affixed to the skull with dental cement and several bone screws. A small metal plate with a hole for a screw soldered to a wire served as the ground. Tetrodes were lowered in small increments after surgery and for 5-7 subsequent days until they reached the target site in dorsal lateral striatum (DV 3.6 - 5.0 mm). Once behavioral training began, the tetrodes were moved as little as possible. However, prior to each training session, tetrodes lacking single-unit activity were incrementally adjusted as necessary to maximize the number of units recorded by each tetrode.

Behavioral Training. For 5-7 days after surgery the tetrodes were lowered further into the brain until they reached the dorsal lateral striatum. The maze consisted of a long starting arm (127 ' 7.5 cm) and two shorter goal arms (33' 7.5 cm) made of black Plexiglas. All arms of the maze were elevated off the floor 22 centimeters and

surrounded at a distance of 14.5-16.5 cm by black walls 41 cm in height. Fresh chocolate was periodically sprinked on the floor of the maze to mask any possible olfaction cues. Rats began each trial in a 20 cm starting platform behind a gate. The gate could be opened by the experimenter to allow rats to traverse the maze. Magnetic plexiglas plates (2.8 ´ 6 cm) with a circular well (diameter: 2.5 cm) were placed at the end of each goal arm for reward delivery. All training occurred in dim red light.

For each trial, rats began on the start platform with the start gate closed. Two seconds of baseline activity was recorded prior to the onset of instruction cues. While still in the start block, rats were presented with one of two tone cues (1 kHz or 8kHz, ~80 dB; 500 ms duration), which indicated whether the right or left goal was baited with chocolate sprinkles. Tones played from a speaker located behind the choice point of the maze. 200-300 milliseconds after the tone ended, the start gate was lowered by the experimenter. The rats could then proceed down the long arm of the maze and choose to turn down either the left or right arm. If the correct arm was chosen, rats were rewarded with chocolate sprinkles placed in the delivery food well. Neuronal recordings were terminated 0.5-1 second after goal reaching. Rats were allowed to finish eating the reward before being guided animals back to the start gate by the experimenter. If the incorrect arm was chosen rats were not allowed to visit the correct arm before being guided back to the start gate. The few trials in which the rat failed to exit the starting block were not analyzed.

Tone-goal associations were counterbalanced between animals. Tones were presented in a pseudo-random order in which the same tone could be repeated up to 3 times in a row. For each session rats received 20 trials of each tone type. Sessions generally lasted

1-2 hours. Training consisted of one session per day, ca 40 trials per session, for up to 57 days. Rats were trained 5-6 days a week, for up to 57 sessions until implants or recordings failed. Rats were then deeply anesthetized (Nembutal, 50-100 mg/kg) and then perfused with 4% paraformaldehyde in 0.1 M phosphate buffer, and 30 μm thick brain sections were stained for Nissl substance to identify recording tracks.

Data Acquisition. Neuralynx data acquisition hardware and cheetah acquisition software (Neuralynx, MT) along with a Med-PC behavioral system (Med Associates, St. Albans, VT) were used for behavioral and neuronal data acquisition. Photobeams supplied timestamps for the following task events: gate opening, out of start, mid-maze run, turn onset, turn offset, and goal arrival. The position of the rat was recorded by a Neuralynx video tracker system that was supplied video images from an overhead CCD camera (60 Hz frame rate). The tracker detected an LED light-source mounted to the back of the headstage. Recorded single-unit activity was amplified (gain: 200-10000) and sampled at 32 kHz once a user-determined threshold was reached. A quiet channel of one of the tetrodes was used as reference.

Behavioral Data Analysis. Run times were calculated as the time it took the animal to go from the out of start photobeam marker to the goal photobeam marker. Reaction time was defined as the time it took the animal to reach out of start photobeam after gate opening. Training-related changes in behavioral accuracy, running times, and reaction time latency were analyzed using ANOVAs. Video tracker data and VHS tapes were reviewed to look for changes in behavior as training progressed. No obvious changes (in

sniffing, lateral movement, etc.) were found.

Learning on this task was defined in two ways. First, we defined a learning criterion of greater than 72.5% correct as performance at this level is statistically greater than chance performance ($p < 0.01$, Chi-square). Second, code written by A. Smith was used to determine which session each rats' probability of correct performance reached and remained above chance (Smith, Frank et al. 2004).

A staging procedure was used following prior work (Barnes, Kubota et al. 2005) to compare data across animals despite differences in behavioral performance and duration of training. Stage 1 corresponded to the first day of training, stage 2 to day 2. Stage 3 was the first day in which the animals performed above 60% correct; stage 4 was the first day animals of greater than 65% correct; and stage 5 was the first two training days in which rats earned 72.5% or more reward. Stages (6 – 9) were subsequent training sessions in which the rats again earned 72.5% or more reward. Note that unless the animals maintained a percent correct performance above 72.5 percent, stages 5-9 are not consecutive. The same day was not used in multiple stages; for example, if an animal performed above 60% on day 1 or 2, the next day in which the animal scored 60% or higher was used for stage 3.

Neuronal Data Analysis. Single units were sorted using Autocut (DataWave Technologies) software. The quality of sorted units was determined by analyzing autocorrelograms and overlays of spike waveforms. Units were analyzed if they had at least 100 spikes in a session. Units were then classified as putative medium spiny neurons based on waveform properties, autocorrelograms, inter-spike intervals, and firing

rates (Tepper, Koós et al. 2004; Barnes, Kubota et al. 2005).

To determine whether each cell was task responsive, a peri-event histogram was generated for 520 millisecond epoch centered around each task event (tone on, 1kHz tone on, 8kHz tone on, gate opening, locomotion onset, mid run, turn onset, turn offset, right turn offset, left turn offset, goal reaching, right goal reaching, left goal reaching, and goal reaching). If the spike count in 4 consecutive 20 msec bins was at least 2 standard deviations above a pre-trial baseline period (1900 to 1400 msec before tone onset) and each of those bins contained at least 2 spikes, the cell was categorized as responsive to that event. Throughout learning, on average, 57% of cells were task-responsive according to this criterion. Cells that did not meet this criterion were considered task-suppressed. Fifty-three percent of cells recorded in stage 1 were task responsive while 56% of cells recorded in stage 9 were task responsive. This was not a significant increase.

To analyze the firing of recorded neurons, the firing of each unit was first normalized with respect to its baseline. The baseline period was defined as 1900 to 1400 msec prior to tone onset, and firing rates during this period were averaged in 10 msec bins over all 40 trials in a session. The baseline firing rate was then defined as the mean firing rate across all bins in this baseline Next, the firing rate during 800 msec time windows centered on each task event was averaged in 10 msec bins over all 40 trials and the baseline firing rate was subtracted from each bin to obtain event-related firing patterns for each unit. Binned data across all events were then smoothed using a five point moving average.

To create population activity plots, the event-related firing patterns were averaged over all task-responsive putative medium spiny cells recorded for each stage of learning.

Because the rat ran each trial at a different speed, 800 millisecond epochs centered on each task event were analyzed. The size of this epoch was determined by maximizing the size of the time window while minimizing the overlap. All major analyses were repeated using non-normalized firing rates, as well as firing rates normalized by the mean over all bins rather than only the baseline period. To determine whether the population activity of task responsive cells changed significantly as training progressed we performed a linear regression for each 10 msec bin across the 9 stages of learning, and computed 95% confidence levels. A significant change was identified for a bin if the confidence bounds of the slope did not overlap zero.

To find the trial epoch that contained each cell's maximum firing rate, each cell's firing was summed over all 40 trials for 800 millisecond time windows around each task event. Data was not smoothed. The 100 millisecond bin containing the maximum firing rate was determined. If two bins had the same maximum firing rate (which happened less than 1% of the time), the cell was counted as having a maximum in both locations. The same procedure was then repeated for 25% of the maximum firing rate.

To find the trial epoch that contained each cell's maximum firing rate, each cell's firing was summed over all 40 trials for 800 msec time windows around each task event. Data was not smoothed. The 100 msec bin containing the maximum firing rate was determined . If two bins had the same maximum firing rate (which happened less than 1% of the time), the cell was counted as having a maximum in both locations. The same procedure was then repeated for 25% of the maximum firing rate.

To determine the consistency of the firing pattern of cells in the dorsolateral striatum from trial to trial across learning, 45 neurons were randomly selected from each of 9

stages of learning. Data were shuffled in 10 msec bins and then we bootstrap the trials of both the real and shuffled data set 1000 times. For each bootstrap the time during the trial in which each neuron reached 86% of its peak firing was ranked in relation to the other neurons in that stage of learning. The standard deviation of both the resulting rank and the bin number in which 86% of the firing occurred for each bootstrapped was determined. Shuffled data received a standard deviation of rank lower than 8 less than 1% percent of the time.

To determine if short term memory was present in the striatum we first took the average firing rate for each cell during a baseline period (-1.9 to -.1.4 seconds before tone onset).

References

Atallah, H. E., D. Lopez-Paniagua, et al. (2007). "Separate neural substrates for skill learning and performance in the ventral and dorsal striatum." Nat Neurosci **10**(1): 126-131.

Bailey, K. R. and R. G. Mair (2006). "The Role of Striatum in Initiation and Execution of Learned Action Sequences in Rats." J. Neurosci. **26**(3): 1016-1025.

Barnes, T. D., Y. Kubota, et al. (2005). "Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories." Nature **437**(7062): 1158-1161.

Barto, A. G. (1995). Models of Information Processing in the Basal Ganglia.

Belin, D., S. Jonkman, et al. (2009). "Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction." Behavioural Brain Research **199**(1): 89-102.

Graybiel, A. M. (1998). "The Basal Ganglia and Chunking of Action Repertoires." Neurobiology of Learning and Memory **70**(1-2): 119-136.

Jog, M. S., Y. Kubota, et al. (1999). "Building Neural Representations of Habits." Science **286**(5445): 1745-1749.

Kao, M. H., A. J. Doupe, et al. (2005). "Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song." Nature **433**(7026): 638-643.

McGeorge, A. J. and R. L. M. Faull (1989). "The organization of the projection from the cerebral cortex to the striatum in the rat." Neuroscience **29**(3): 503-537.

Olveczky, B. P., A. S. Andalman, et al. (2005). "Vocal experimentation in the juvenile songbird requires a basal ganglia circuit." PLoS Biol. **3**: e153-e153.

Packard, M. G. and J. L. McGaugh (1996). "Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning." Neurobiology of Learning and Memory **65**(1): 65-72.

Rosenbaum, D. A., S. B. Kenny, et al. (1983). "Hierarchical control of rapid movement sequences." Journal of Experimental Psychology: Human Perception and Performance **9**(1): 86-102.

Sakai, K., K. Kitaguchi, et al. (2003). "Chunking during human visuomotor sequence learning." Experimental Brain Research **152**(2): 229-242.

Schmitzer-Torbert, N. C. and A. D. Redish (2008). "Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype." Neuroscience **153**(2): 349-360.

Schultz, W., P. Dayan, et al. (1997). "A Neural Substrate of Prediction and Reward." Science **275**(5306): 1593-1599.

Smith, A. C., L. M. Frank, et al. (2004). "Dynamic Analysis of Learning in Behavioral Experiments." J. Neurosci. **24**(2): 447-461.

Tepper, J. M., T. Koós, et al. (2004). "GABAergic microcircuits in the neostriatum." Trends in Neurosciences **27**(11): 662-669.

Wilson, M. and B. McNaughton (1993). "Dynamics of the hippocampal ensemble code for space." Science **261**(5124): 1055-1058.

Yin, H. H. and B. J. Knowlton (2006). "The role of the basal ganglia in habit formation." Nat Rev Neurosci **7**(6): 464-476.

Figures

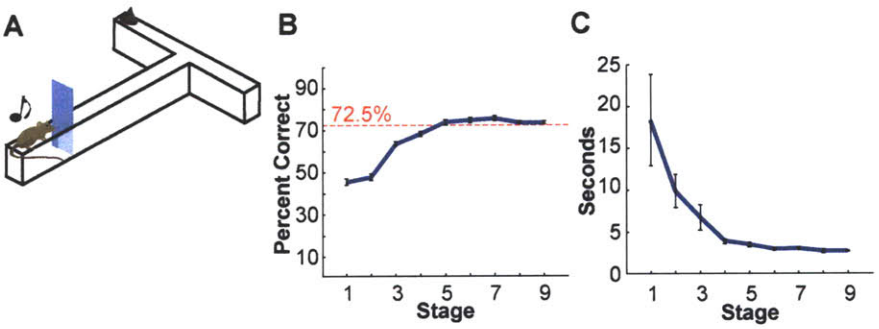


Figure 1. The T maze task, design and performance. ***A,*** Rats first heard a 500 ms tone identifying which arm would contain chocolate reward. After the tone terminated, a gate opened to allow the animal to proceed down the runway and select one of the two arms. ***B,*** Average percent correct by stage of all animals across learning stages showing that animals learn to our criteria of 72.5% correct ($p<.0001$). ***C,*** Average run time by stage of all animals throughout learning showing that animals ran the task faster as training progressed ($p<.001$). Error bars indicate standard error of the mean.

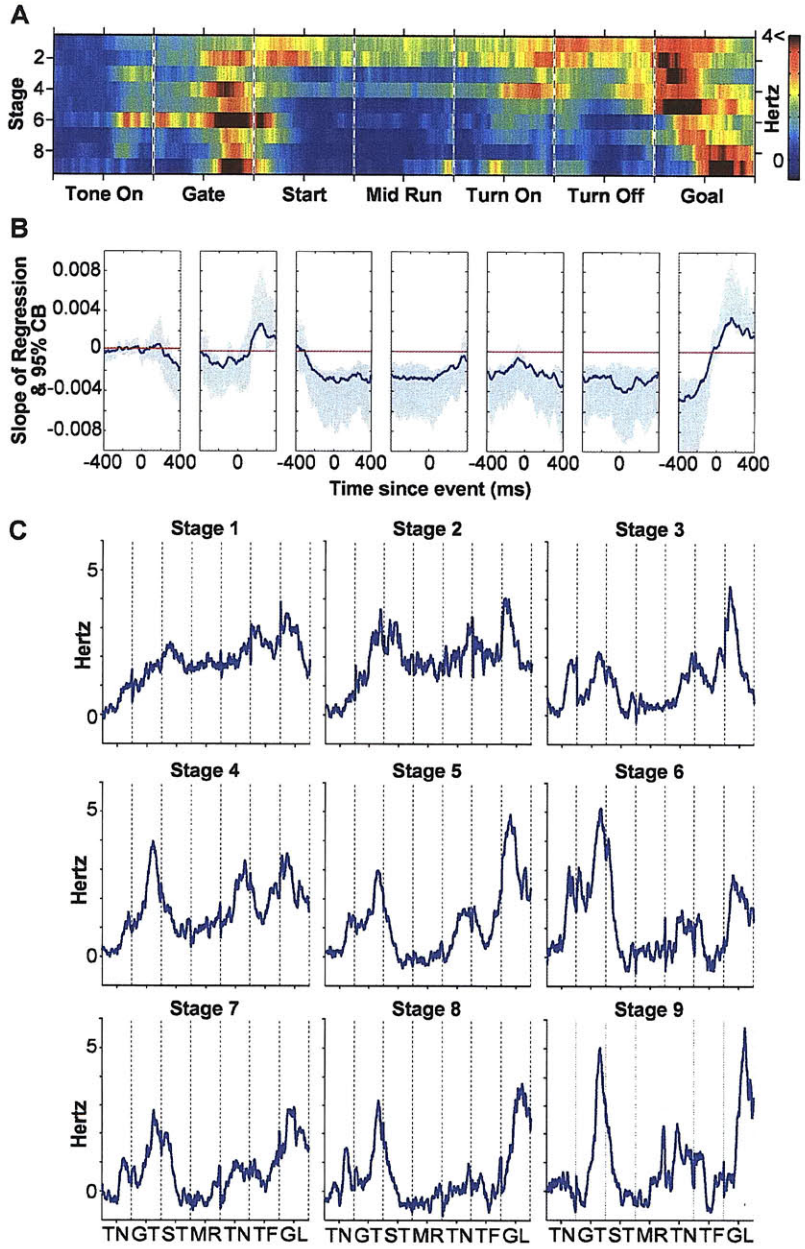


Figure 2. Average normalized per trial spike rates across training. ***A,*** Task responsive medium spiny neurons. Panels show 800 ms period of cell activity around baseline, tone, gate, start, mid run, turn onset, turn offset and goal events across learning stages. Each 10 millisecond bin was normalized by subtracting an average pre trial firing rate; plotted on a pseudo color scale of average firing rate. Note that task responsive cells show more activity at the beginning and end of trial runs as training progresses. ***B,*** The slope of the regression for each 10 millisecond bin through learning with 95% confidence bounds showing which bins significantly increased (above the red line), or decreased (below the red line) as training progressed. Data shown for 800 milliseconds panels for each event ***C,*** Same data as in A, plotted in line format to show the progression of the firing pattern throughout learning.

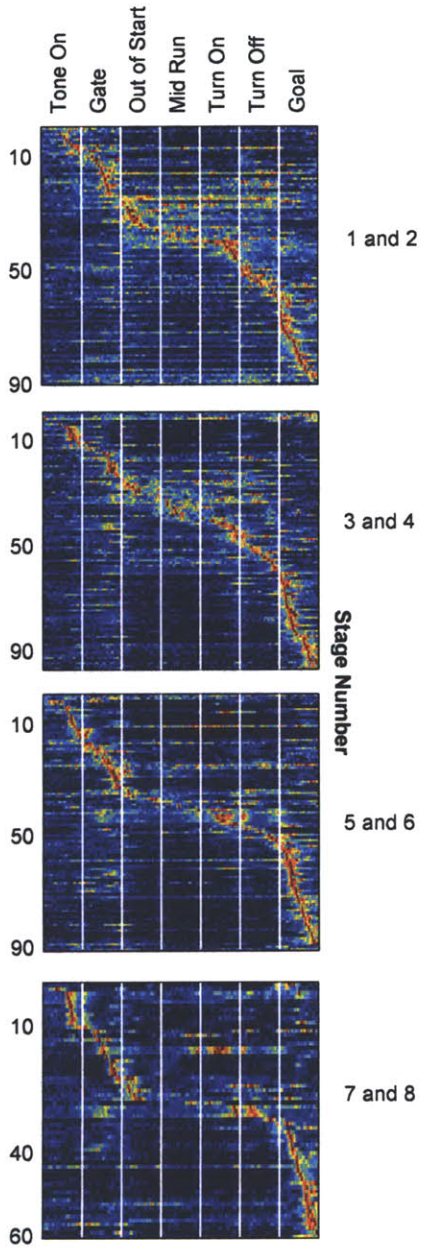


Figure 3. Firing Pattern of Single Cells. ***A,*** task responsive firing rate plotted for 800 ms around each task event in 10 ms bins. The data were smoothed with an 8 pass filter and than min-maxed. Each cells maximum firing rate is plotted in red. Cells sorted base on when they first hit 86% of their maximum firing rate. Each horizontal line represents the normalized firing rate for one cell as the trial progressed.

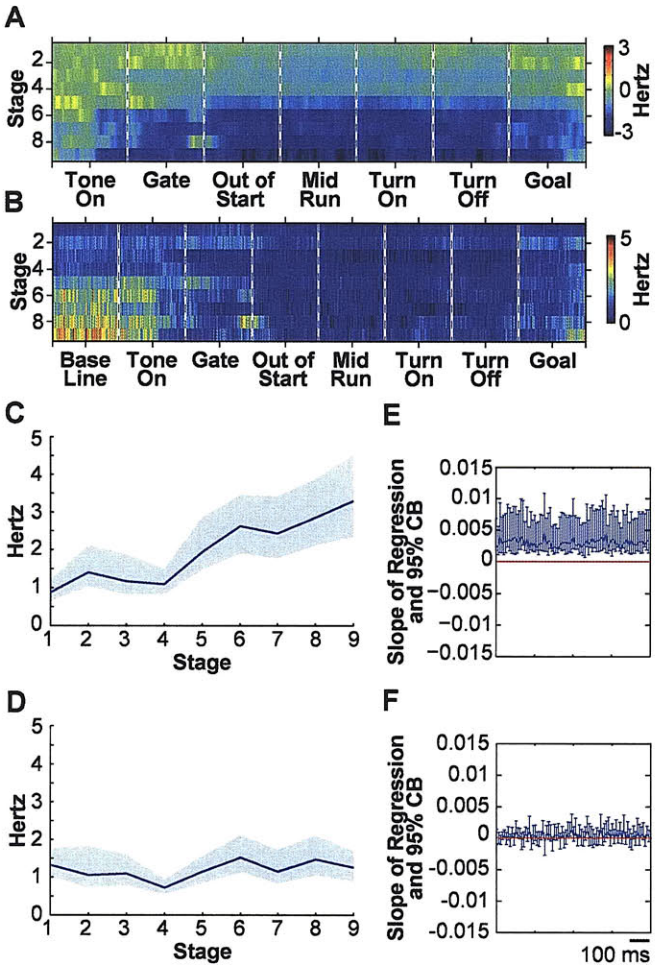


Figure 4. Baseline firing rates. ***A,*** Normalized average firing rate of non-task responsive putative medium spiny cells plotted in 10 ms bins during 800 ms time windows around task events as training progresses. ***B,*** Raw average firing rate of non-task responsive putative medium spiny cells plotted in 10 ms bins during 800 ms time windows around task events as training progresses ***C,*** The baseline firing rate for non-task responsive and ***D,*** task responsive cells throughout learning. Shaded area indicates 95% confidence bounds. ***E,*** The slope of the regression for each 10 millisecond bin through learning with 95% confidence bounds showing which bins significantly increased (above the red line, or decreased (below the red line, as training progressed for non-task responsive cells and ***F,*** task responsive cells. Note that the baseline significantly increased for non task responsive cells.

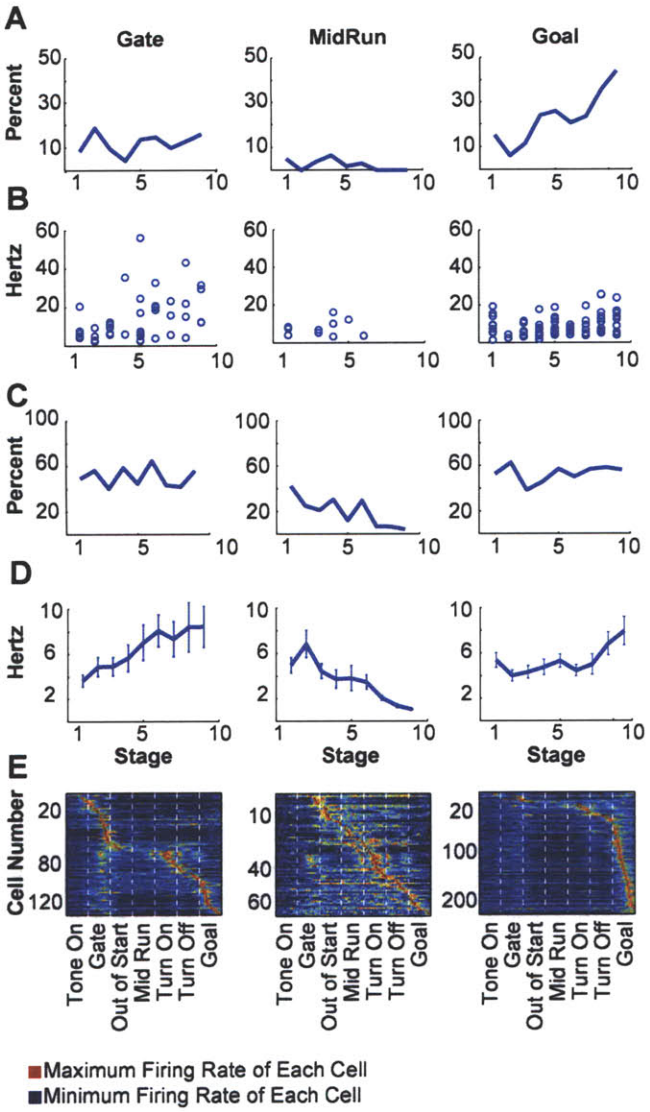


Figure 5. Maximum firing rate across learning. ***A,*** Percentage of tasks responsive cells hitting their maximum firing rate during 400 milliseconds after gate, 400 milliseconds during mid run, and 400 ms after goal. The percentage of cells increases for goal cells ($p<.01$) ***B,*** The size of the maximum firing rate for those cells that hit their maximum firing rate during each event. ***C,*** Percentage of tasks responsive cells hitting at least 25% of their maximum firing rate during 400 milliseconds after gate, 400 milliseconds during mid run, 400 ms during turn and 400 ms after goal. ***D,*** The size of the maximum firing rate for cells that hit 25% of their maximum firing rate during each event. This significantly increased for cells responding to gate ($p<.01$). Error bar indicate the standard error if the mean. ***E,*** Responses of subpopulations of cells. The raw session firing rate of each cell is plotted for 800 ms around each task event in 10 ms bins. Task responsive cells from all stages of learning are plotted. The data were smoothed with an 8 pass filter and than min-maxed. Maximum firing rate of each cell is plotted in red. Cells are sorted based on when they first hit their maximum firing rate. Panels show cells that increased their firing rate in response to gate, mid run and goal.

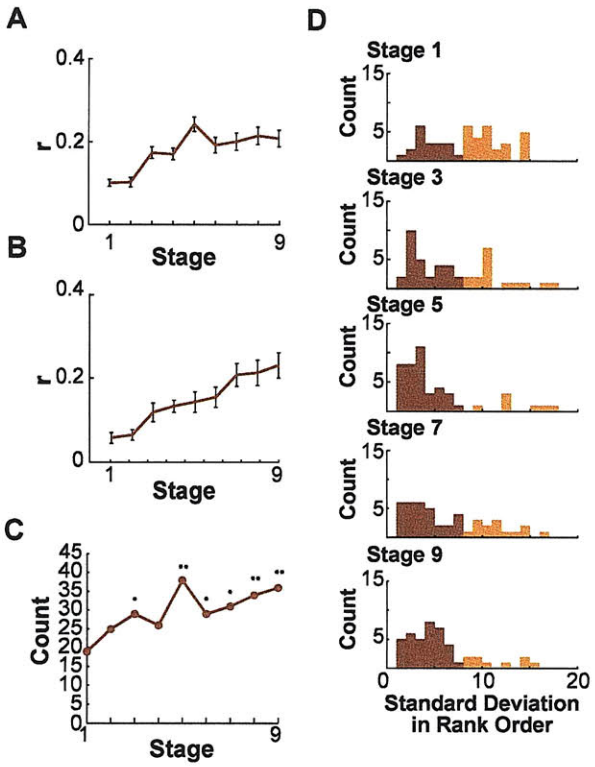


Figure 6. Consistency of firing patterns throughout learning. ***A,*** As training progresses the amount of cells with a standard deviation of 8 or lower increase. ***B,*** Average correlation of trial firing pattern increases as training progresses for 800 ms time period centered on gate opening. ***C,*** Average correlation of trial firing pattern increases as training progresses. ***D,*** A sum over 40 bootstrapped trials was taken 1000 times. The cells were then ranked based on when they hit their maximum firing rate. The standard deviations of the resulting ranks are plotted for odd stages. Dark orange bars indicate cells whose standard deviation in rank was lower than 8.

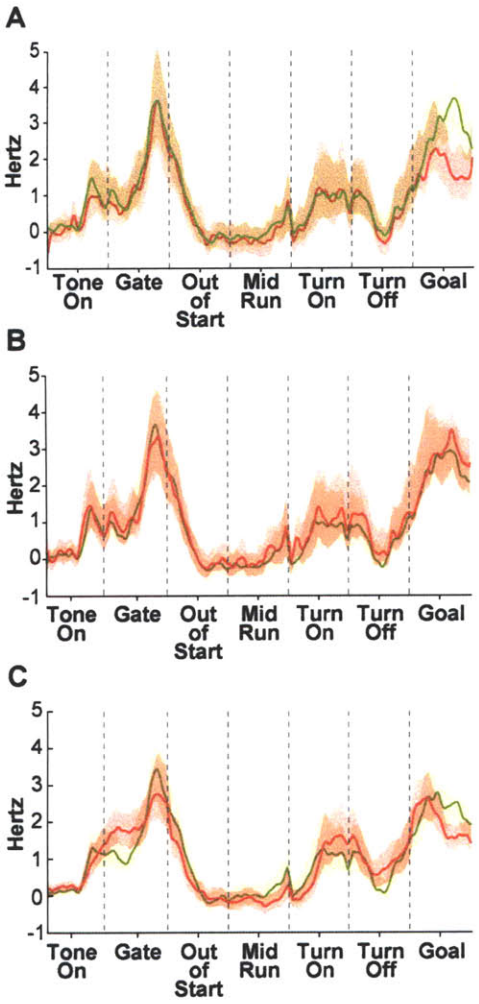


Figure 7. Average firing rate for different types of trials throughout learning. ***A,*** Normalized average firing rate of task responsive putative medium spiny cells recorded in stage 5 to 9 plotted in 10 ms bins for correct (green) and incorrect (red, trials during 800 ms time windows around task events as training progresses. Shaded area indicates 95% confidence intervals. ***B,*** Normalized average firing rate of same cells as in A, for trials following a correct trial (green) and trials following an incorrect trial (red). Note that these firing patterns are similar. ***C,*** Normalized average firing rate of task responsive putative medium spiny cells recorded in stage 5 to 9 plotted in 10 ms bins for sessions wherein rats responded or 60% of the time correctly or less (red, top n=293) and sessions where rats responded or 70% of the time correctly or more (green, top n=240).

1 columns wide Earlytone Figure 8 11/16/2009

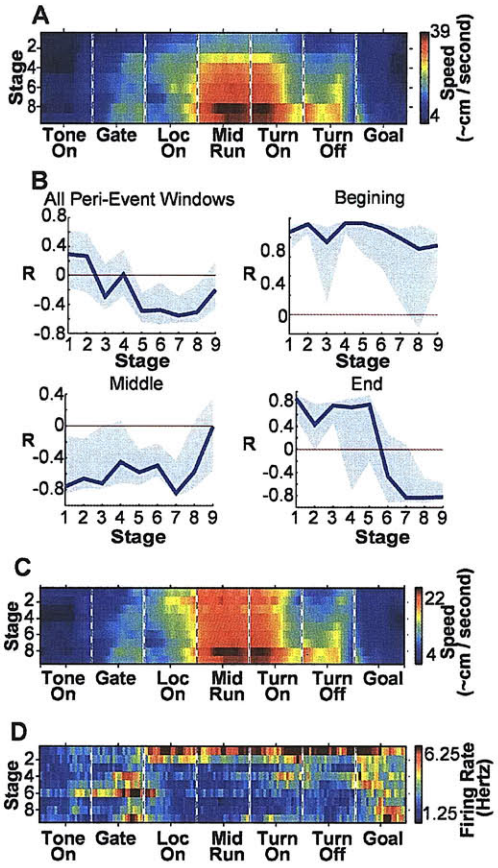


Figure 8. Comparison of Firing Rate and Speed of Rats . ***A***, Average speed as training progresses. The speed for each animal was calculated and multiplied by the amount of task responsive cells contributed by that animal for that stage. ***B,*** Correlation of average speed and average firing rate by stage for all 800 millisecond epochs, tone and gate epochs, out of start, mid run, turn on, and turn off epochs and goal epoch. Shaded areas represent 95% confidence intervals. ***C,*** Plot of speed for trials in which speed was similar throughout learning. For each stage, 246 trials were randomly selected from those trials whose speed for the 800 ms around mid run was at least 1 standard deviation below the mean speed around mid run for stages five through eight. Other time periods were also used to determine if speed was similar and the results were similar. ***D,*** Plot of firing rate for included trials demonstrating the changes in firing rate as training progresses still exist.

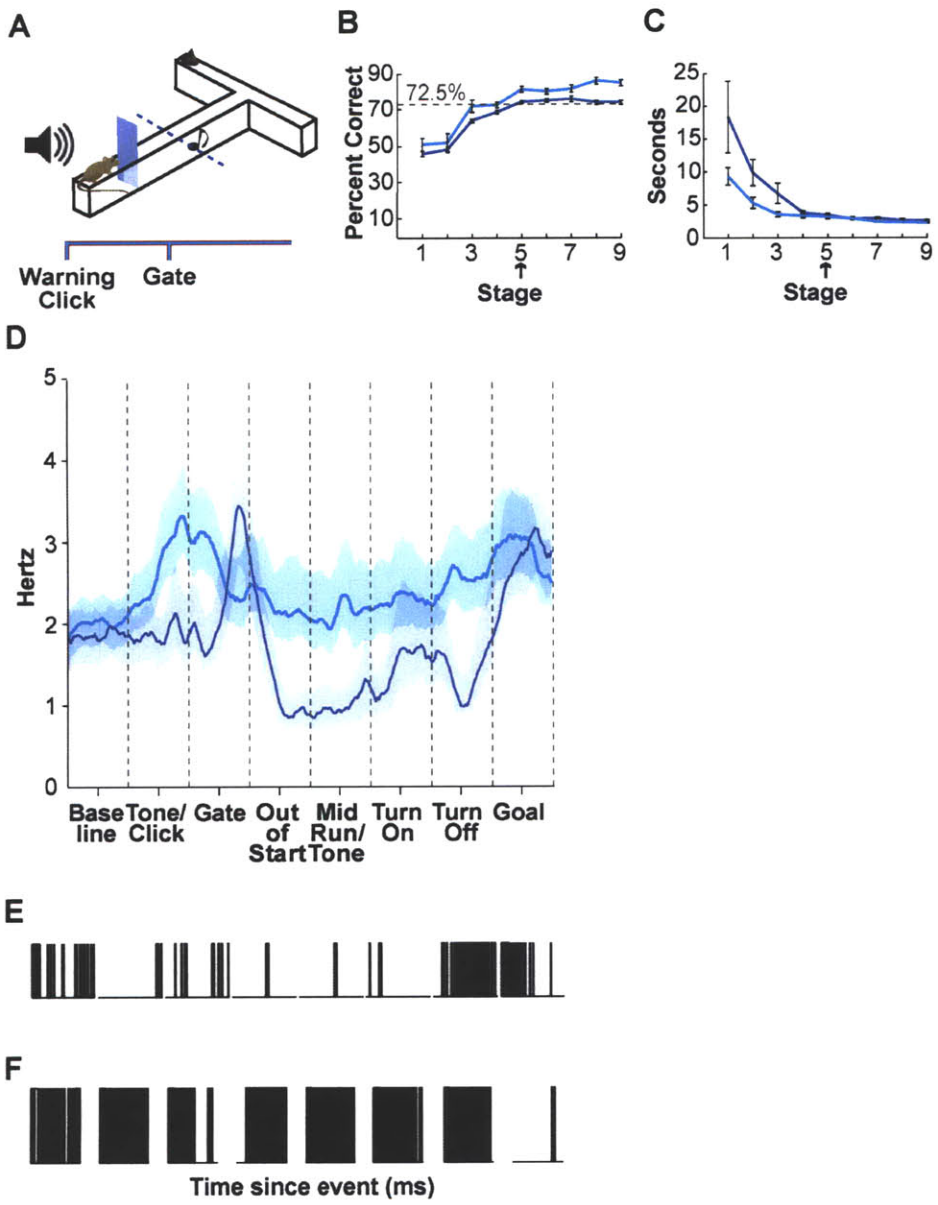
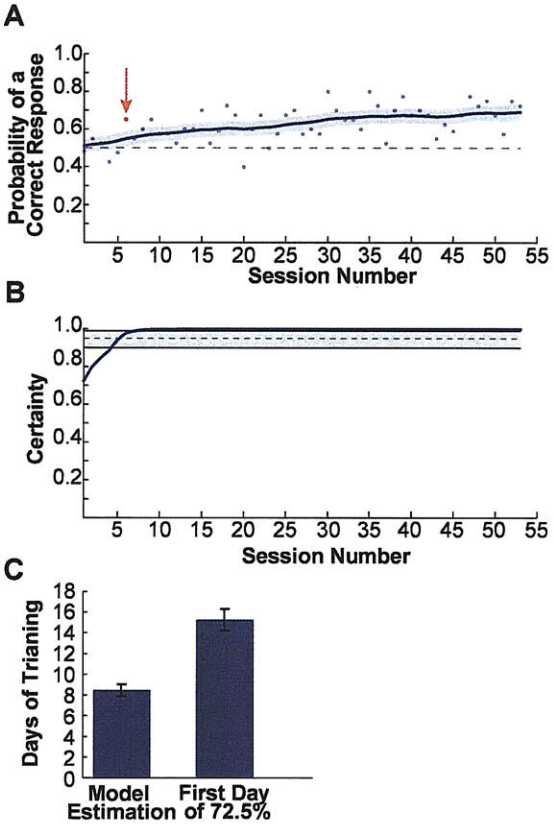
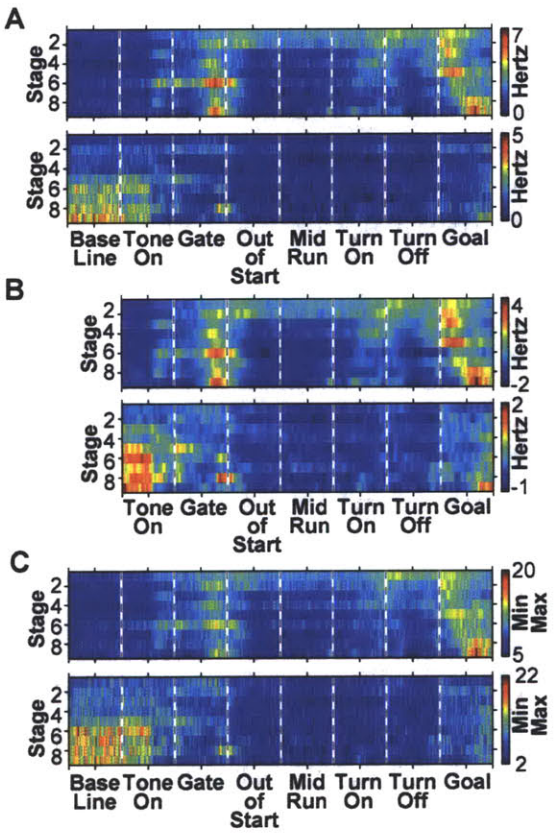


Figure 9. Average normalized raw firing rate of task responsive cells recorded during overtraining sessions in two tasks. ***A,*** In the previous task rats first heard an auditory click that only signaled the lowering of the drawbridge gate. The tone identifying which arm would contain chocolate reward was turned on mid run and remained on until rats reached one of the two goals. ***B,*** Average percent correct of both groups of rats (n=7, blue, (n=7, red, across learning stages showing that both animals learn to the criteria of 72.5% correct. ***C,*** Average run time by stage of all animals throughout learning showing that animals learned to run this task faster as training progressed. Error bars indicate standard error of the mean. ***D,*** Data from this task (blue), and another version (aqua), where the tone turned on mid run. Note that in our task, the increase in firing rate that occurs at the beginning of the task occurs after the gate and not the tone, also note that in this task (blue, the firing during the middle of the task is decreased. Shaded bars indicate 95% confidence intervals. ***E,*** P values per bin for differences between the two tasks Black bars indicate p values less than .05. ***F,*** P values per bin for differences in slopes across learning per bin between tasks.

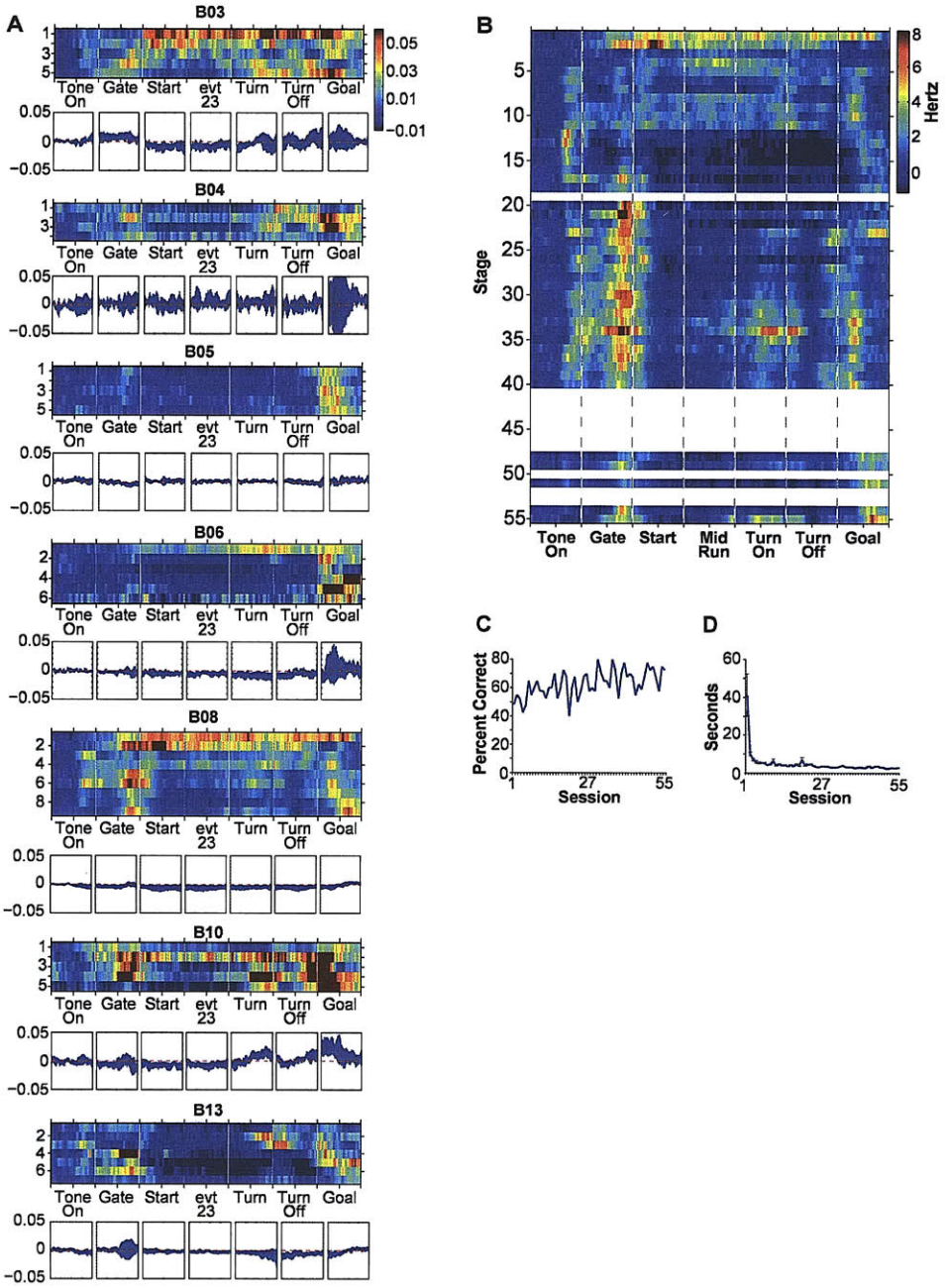
Supplemental Figures



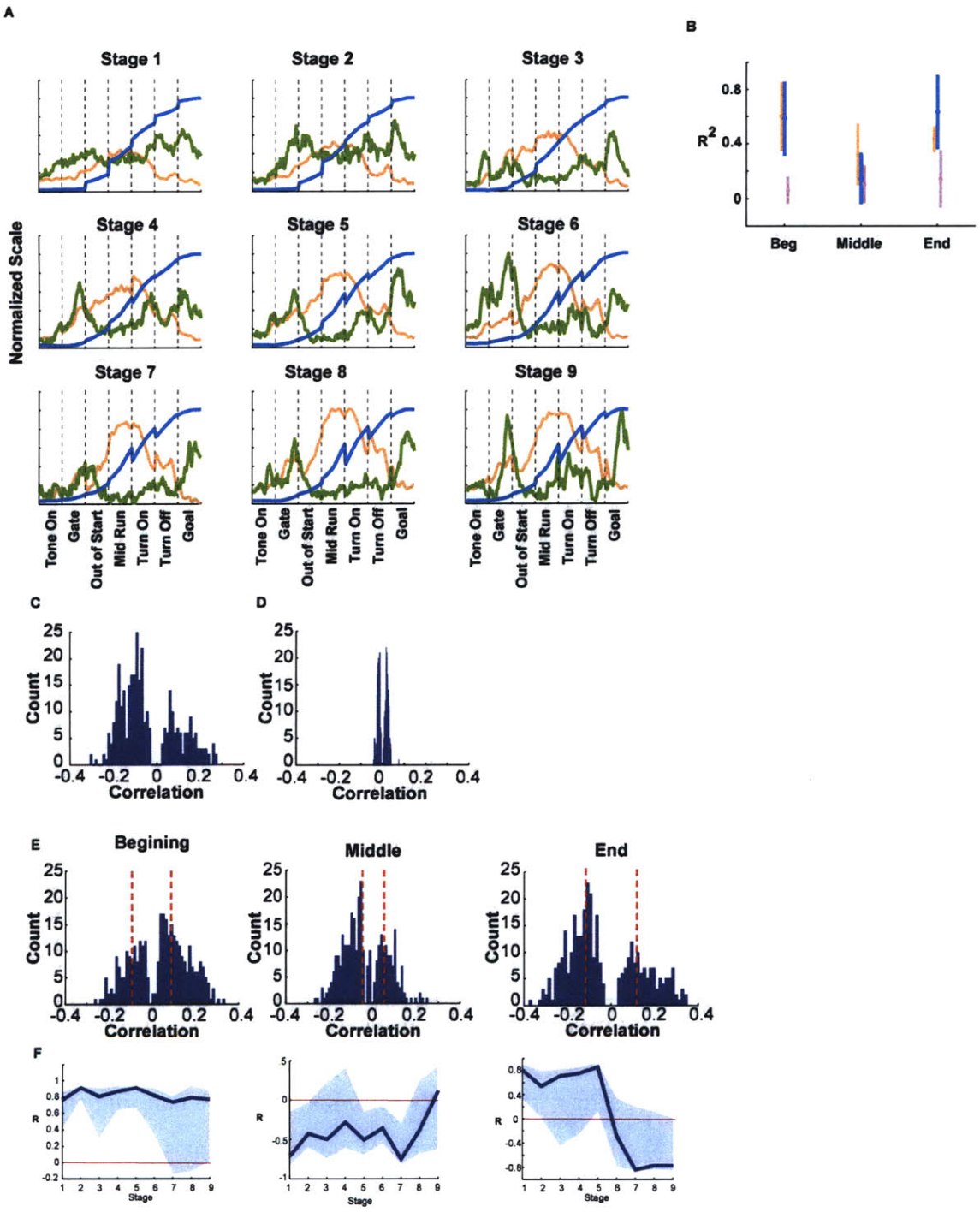
Supplemental Figure 1. Model of learning. ***A,*** The probability of a correct response for one animal. Dashed line indicates chance performance. Red dot indicates the day of learning according to the state space model by Smith et al. ***B,*** Certainty of above chance performance by session. Shaded gray area indicates confidence intervals of certainty ***C,*** Average days until learning with the model by Anne Smith compared to a criteria of 72.5% ($p < .01$). Error bars indicate standard error of the means.



Supplemental Figure 2. The progressive patterning of the firing pattern in the dorsal lateral striatum is apparent regardless of how firing rates are calculated. **A,** Raw average firing rate of task responsive (top) and non task responsive (bottom) putative medium spiny cells plotted in 10 ms bins during 800 ms time windows around task events as training progresses. **B,** Same data as in A, except data were normalized for each cell by subtracting the firing rate of the average bin for that cell. **C,** Same data as in A, but before averaging, each cell was placed on a 1 to 64 scale. Note that for task responsive cells, all methods of calculating firing rate yield similar results. For non- task responsive the baseline increases as training progresses.



Supplemental Figure 3. Individual behavior and neuronal firing pattern of individual rats. ***A,*** The normalized per stage firing rate for the cells of the individual animals. ***B***, Normalized per session firing rate of task responsive cells of a rat showing a clear accentuation of the task beginning and end emerging over training. Data shown for those sessions in which at least 5 cells were recorded (gray=too few cells). Data plotted for 800 millisecond windows around each task event in 10 ms bins and normalized by subtracting a pre-tone baseline. ***C,*** The percent correct turns by this rat per session. ***D,*** Run time of this rat by session.



Supplemental Figure 4. Speed and Firing Rate. ***A,*** Normalized Firing rate (green) and speed (orange) by stage. ***B,*** Histograms of adjusted R squared values for regression with firing rate of cells in the dorsal lateral striatum. 336 single putative medium spiny cells from all stages of learning were used. Data were first averaged of all trials in a session.

Chapter Four

The Relationship between the Firing Rate in the Sensorimotor Striatum and Motor Movement

Abstract

The sensorimotor striatum receives a substantial amount of lower level motor input as well as higher level limbic input. This structure has been found to show training induced activity in a variety of tasks. However, in these tasks there is both a cognitive component (such as learning to associate a specific tone with a specific direction of turn) as well as motor component (such as running faster as training progresses). In order to understand which aspect(s) of learning are responsible for the training-induced changes seen in our T-maze task, we have performed several single cell and ensemble analyses. In this task, rats must learn to turn left or right in response to tones that predict the location of a chocolate reward. We find that between 5 and 24% (depending on whether the state of the rat is evaluated) of cells are sufficiently correlated with the running speed of the animal to account for at least 5% of the variability in firing rate. Excluding these well correlated cells does not destroy the training-induced pattern found. The pattern of training-induced activity in the firing patterns of sensorimotor cells remains when the running speed of animals is held constant. We found similar levels of correlation for speed and place as these variables themselves are well correlated in this task. Furthermore, in a control task in which rats' running speeds also increased, but no association between the tone played and the direction of reward was present, the training-induced pattern was not found. We conclude that speed and/or position play an important role in the activity that develops in the dorsolateral striatum, while acceleration appears to play a lesser role.

Introduction

The dorsolateral striatum, often called the sensorimotor striatum, receives input from several sensory and motor structures. It also receives input from cognitive areas such as the substantia nigra pars compact and mesocortex (McGeorge and Faull 1989). Here we examine what type and to what degree purely motoric striatal responses should be expected and are found in recordings of the dorsolateral striatum, and to what extent changes in movement may account for re-patterning of the dorsolateral striatal neural activity as habits are learned.

Trytek et al. estimated that 80-90% of striatal cells change their firing rate with movement based on open field recordings but it is unclear what aspects of movement trigger activation of these cells(Trytek, White et al. 1996). Aldridge and Berridge have extensively studied striatal responses during grooming, an innate behavior in rats. Rodent grooming has syntactic (rule-driven) sequences and more random movement patterns (although not completely lacking in syntactic structure) (Aldridge 2004). They found that activity of 41% of cells in the striatum coded the sequential pattern of syntactic chains (movement in a specific cognitive context) while only 14% coded purely motoric properties of grooming movements by firing consistently each time a particular movement was made (Aldridge and Berridge 1998). West et al. also looked at the percentage of cells in the dorsolateral striatum that respond to movement parameters and found that 18% of cells were related to movement of a specific body part (West, Carelli et al. 1990; Carelli, Wolske et al. 1997). Furthermore, they found that these seemingly

pure movement-related responses faded with training, further indicating the context dependency of movement related activity in the dorsolateral striatum (Carelli, Wolske et al. 1997). A wide array of experiments have shown that cells in the striatum respond to cognitive features during movement such as action value and reward (Apicella, Ljungberg et al. 1991; Schultz and Romo 1992; Apicella, Legallet et al. 1997; Tremblay, Hollerman et al. 1998; Jog, Kubota et al. 1999; Packard and Knowlton 2002). The majority of cells in the striatum are not activated by simple motor movements per se, but rather specific movements in the context of a particular sequence, or particular cognitive process.

Procedural tasks, such as the T-maze, have a motor component which can change in parallel with learning and habit formation. For example, rats run faster, more consistently, and with more stereotyped paths through the maze as the task becomes well learned and habits are formed. Therefore, when analyzing data from procedural learning tasks it is important to carefully analyze the contributions of movement parameters (e.g., speed, acceleration) as well as the contribution of learning to neural activity results.

In this task speed was measured by tracking a LED mounted on the back of the animal's headstage. The location of the LED was tracked using Neuralynx video tracker software and was sampled at 60 Hertz. Data were analyzed with a 40 ms bin size or a 1 centimeter grain. This task was not designed to answer questions relating to the specific muscles used by the animals. Consequently, we do not have an optimal data set for these measures. We used one LED which provided a two-dimensional record of the animal's head location, but we could not measure head direction or movement occurring in the depth plane. Despite these design limitations, we were able to analyze the movement

variable that perhaps changes most dramatically as training progressed: speed. We were also able to analyze one that fluctuates repeatedly during the task: acceleration.

Figure 1 shows the striking changes in the normalized average firing rate of cells in the dorsolateral striatum as rats learn a discriminative T-maze task. It also shows the normalized average speed of rats as they perform this task. Here we examined the relationship between firing rate and speed, acceleration, and position.

Results

Single cell, trial by trial

In order to determine whether the patterning seen to develop in striatal firing in our task (i.e., accentuation of firing of task-responsive neurons at the beginning and end of the maze) is due to changes seen in the speed of the animal as training progresses, we examined the correlation of firing rate and speed for each task-responsive cell. We first determined the optimal lag time for each cell. The correlation between speed and firing rate was calculated for each trial for each cell by shifting the speed in time for the range of 12 bins (480 ms) in a positive and negative direction around zero (zero defined as the simultaneity of the speed and firing rate measures). The highest mean correlation was then determined for each cell. **Figure 2a** shows the distribution of mean correlations between firing rate and speed for each task responsive cell (n=336) at its best lag. The same lag calculation procedure was followed for a data set in which the firing rates were shuffled. **Figure 2b** shows the same plot for these shuffled data. Less than 1 percent of the shuffled data's highest mean correlations exceeded R=.05.

A large number of cells showed a significant correlation with speed compared to the shuffled data. The magnitude of the correlation, for the majority of cells, was modest. 304/336 cells showed a significant correlation (larger in magnitude than .05 which occurred less than 1% of the time in shuffled data). The range of the r value correlations was -0.31 to 0.27. The mean for all task responsive cells was R =-0.04 (**Figure 2a**).

Of task responsive cells examined, 99/336 (29% mean R= 0.14) showed a significant positive correlation (higher than .05), while 205/336 (61%, mean R= -0.13) showed a significant negative correlation (lower than -.05). There were significantly more cells showing a negative correlation than a positive one (205 compared to 99 $p<.001$). Animals hit the maximum speed in the middle of the maze and were slower at the ends (**Figure 5a**). Therefore, cells that fired more at both the beginning and end of the maze would tend to show a negative correlation with speed.

Figure 2c shows the optimal lag for each cell plotted against its correlation coefficient. The optimal lag centered around zero, and there were not significantly different amounts of positive and negative lags ($p>.05$). The amount or direction of lag did not predict the direction or magnitude of the correlation. This was also true for the shuffled data (**Figure 2d**).

Thus, speed in this single cell trial by trial analysis accounts for up to 10% of the variability in the activity of the most highly correlated cells, and on average across all cells in this data set accounts for less than 1% of the variability found in firing rates of single cells. 5% of cells had a correlation high enough to account for 5% of the variability in the sample. The range of correlations with speed is similar to that found previously for LFPs recorded in the medial striatum (DeCoteau, Thorn et al. 2007). **Figure 2e**

shows the cell by cell correlations broken down by stage, and illustrates that the correlations did not change systematically during the course of learning.

These low correlations in single cells with speed is in contrast to data previously published by the Mizumori laboratory (Mizumori, Yeshenko et al. 2004; Yeshenko, Guazzelli et al. 2004; Eschenko and Mizumori 2007), in which a subset of cells in the dorsomedial striatum were found to be well correlated with acceleration and/or speed. This may be due to the difference in recording location. Unlike the dorsolateral striatum, the dorsomedial striatum receives input from the hippocampus via the entorhinal cortex (McGeorge and Faull 1989; Antonius B. Mulder 2004).

If the correlation between speed and firing rates of cells in the sensorimotor striatum changes as animals traverse the maze, it may explain why the cell by cell correlations calculated for the entire trial are relatively small (Yeshenko, Guazzelli et al.). Note that in **Figure 1,** both at the beginning of the maze runs (before tone on) and during the mid-run period (in stage 5 for example) the average ensemble firing rate is ca 0 Hz (relative to baseline). However, the speed of the animal at these two time periods is very different: it is lowest at the beginning of the maze runs before tone on, and is highest during the mid-run period. This may because the brains of rats may be in a different state early in the trial before the start gate is lowered than later in the trial when the animal is free to run towards goal. Indeed, theta is present only after rats start to transverse the maze and starts to fade as animals approach goal (DeCoteau, Thorn et al. 2007).

We therefore next analyzed the data in three parts: the beginning (encompassing 800 milliseconds around tone on and gate), the middle (encompassing 800 milliseconds around out of start, mid run, turn on and turn off) and the end (encompassing 800

milliseconds around goal). In the beginning of the trial, we found that more cells were positively correlated than negatively correlated ($p<.005$). The range of correlations was -.27 to .34. Twenty-five out of three hundred and thirty-six (~7%) showed a positive correlation high enough to account for 5% of the variability seen in the firing rate while 2/336 (<1%) showed a negative correlation low enough to account for 5% of the variability seen in the firing rate. Data are shown in **Figure 3**.

During the middle of the trial, we found that more cells were negatively correlated than positively correlated ($p<.005$). The range of correlations was -.27 to .25. 1/336 showed a positive correlation high enough to account for 5% of the variability seen in the firing rate, while 5/336 showed a negative correlation low enough to account for 5% of the variability seen in the firing rate (**Figure 3**).

During the end of the trial, we found that more cells were negatively correlated than were positively correlated ($p<.005$). The range of correlations was -.38 to .35. 33/336 (~10%) showed a positive correlation high enough to account for 5% of the variability seen in the firing rate, while 34/336 (~10%) showed a negative correlation low enough to account for 5% of the variability seen in the firing rate. Of these 33 were classified as goal responsive.

By taking into account the state (beginning, middle, end) from which the data are collected the amount of cells that had a correlation with speed large enough to account for 5% of the variability in firing rate increased from 5% to 24% (79/336). This corresponds well with previously reported numbers for the percentage of cells in the dorsolateral striatum showing motoric responses (West, Carelli et al. 1990; Trytek, White et al. 1996; Aldridge and Berridge 1998; Aldridge 2004).

Single cells averaged across sessions

It could be that the relationship of the firing of cells in the striatum and speed is hard to discern on a trial by trial basis because the spike counts in each trial are low. We therefore did a series of analysis on the data that were first averaged of all 40 trials in the session. Single cell regression analyses were then performed for each cell with speed, acceleration and position together, each factor separately, and speed and position together. Data were first smoothed with a window of 100ms. Eight hundred millisecond time windows around each task event were examined. Data were averaged across all 40 recorded trials and then binned in 40 ms bins. Regression analyses were then performed for each cell separately. A frequency diagram was then constructed of the adjusted R squared values in order to show how well each variable could account for the variability of firing found in the dorsal lateral striatum. Data were analyzed for the entire trial and for the beginning, middle, and end portions.

Speed and position could account for a similar amount of variability (11 to 30%) both in the entire trial and during each portion of beginning, middle, and end. Acceleration accounted for substantially less variability (2% to 9%). More of the variability could be accounted for during the beginning and end portions of the maze than the middle or for the entire trial taken together **(Figure 4 and Table 1)**.

Population coding: average firing rate during stages of learning

We next correlated the average firing rate (**Figure 5a**) and average speed (**Figure 5b**) for all cells recorded in each stage of learning (**Figure 5c**). Early in training, the

correlation was positive (though not significant, see **Figure 5c**) and primarily became negative as training progressed. Therefore, the relationship between speed and firing rate was not consistent. This was also true at a range of different lags (**Figure 5d**).

To further determine whether changes in speed underlie the changes we found in firing rate as training progressed, we examined only those trials in which the animal ran at a similar speed. We determined the average speed and its standard deviations (SDs) for sessions included in stages 5 through 8 (those stages where the average speed and average firing rate were consistently negatively correlated) and we excluded trials in stages 1 though 9 that were lower than one standard deviation below this mean (**Figure 6**). We matched the number of trials in each stage by randomly selecting 246 trials for each session (the lowest number of trials for any session that fit the criteria). Therefore, in this dataset, the speed is similar for all of the trials plotted. Note that the pattern of ensemble firing still changed with learning. This outcome demonstrates that the changes in the firing patterns seen here are dissociable from changes in the speed of the animals; when the speed was held constant, the pattern still emerged over learning.

Using multilinear regression on the averaged firing rate, speed, acceleration and position, we have found that speed can account for 15% (s.e. +/-1.3%) of the variance found in the averaged firing rate of cells in the dorsal lateral striatum. Using position as well as speed, we can account for 36% (s.s.e. +/-.7%) of the variance in the averaged firing rate (see first row of **Table 2**). This analysis was performed on the averaged data collapsed over all learning stages. We also performed this analysis on each stage of learning separately and then took an average for all nine stages (see first row of **Table 3**). Results were similar.

Thus, by these analyses, speed could account for part of the variance in firing rate. We estimate that it accounts for roughly ~15%. By taking into account the state (beginning, middle, end) from which the data are collected, the level of correlations we found increased significantly.

Like the single cell, single trial analysis, we next broke the trial time into beginning, middle, and end of the trials. This increased significantly the level of correlations we found. There is a strong positive correlation between the average firing rate and the average speed of the animals at the beginning of the trial. During the middle of the trial this correlation turns strongly negative. At the end of the trial, the relationship is less consistent but starts strongly positive and turns strongly negative (**Figure 5e**).

Next, excluding cells that, during these three time periods, in the single cell trial by trial analysis were found to be well correlated with speed, we plotted the average firing rate (**Figure 7**). When excluding cells that are well correlated with speed, the accentuation of the beginning and end trials is still present suggesting that the changes we see in the average firing rate can not be exclusively due to changes in speed.

Next, we used multiple linear regression to compare the potential contributions of speed, position, and acceleration to the variance in firing rate during each portion or state of the trial. The results are in **Table 2 and Table 3.**

This analysis was carried out both on data collapsed over all stages of learning (**Table 2**) and for data from each stage of learning which was then averaged (**Table 3).** After collapsing over stages of learning and breaking the trial time into three epochs, I found that 88% (per stage range: 14-88%, mean = 60%) of the variance in firing rate at the beginning of the task can be accounted for by speed. The collapsed data show that

together speed and position can account for a total of 89% of the varaince (per stage range: 45-92%, mean 74%) during the beginning of the task. Forty percent (per stage range: 0-60%, mean 32%) of the variance in firing rate during the middle of the task can be accounted for by speed. Speed and position together can account for 47% of the variance (per stage: 0-69%, mean 40%). During the end of the trial, speed does not account for a significant amount of the variability of firing rate in the combined dataset but in per stage data can account for an average of 43% (range 27-55%). Together, speed and position do not account for a significant amount of the variance compared to chance during the end of the trial in the combined dataset but per stage accounts for an average of 73% (range 50-93%). The percent of variance in firing rate that could be accounted by speed alone was in general similar. Please see **Table 2** and **Table 3** for a complete summary. **Figure 8** shows the variation in the adjusted R squared values during learning and shows that they are similar for position and speed.

Acceleration

We next examined the relationship between the firing rate of the striatal neurons that we recorded from and the acceleration of the animals. We examined the correlation of firing rate and acceleration for each neuron. The optimal lead-lag of acceleration and firing rate was calculated for each cell. **Figure 9a** shows the distribution of correlations with acceleration for all task responsive cells (n=336) at their best lag. **Figure 9b** shows the optimal lag for each cell plotted against the correlation coefficient. The range of the correlations was -0.22 to 0.20. The mean for all task responsive cells was R= -0.02. 234/336 (70%) of cells showed a significant correlation. 93/336 (28% mean R= 0.09)

task responsive cells examined showed a significant positive correlation (i.e. when acceleration increased, firing rate increased) while 141/336 (42% R= -0.09) cells showed a negative correlation (i.e. when acceleration increased, firing rate decreased).

There are significantly more cells showing a negative correlation than a positive one (208 compared to 128 $p<.01$). There were also more optimal lags with changes in firing occurring after changes in speed ($p<.001$) (**Figure 9b**). This can in part be explained by the strong acceleration that occurred as animals started the maze runs being followed (~0 to 500 ms) by the firing rate increase that developed in the striatal firing patterns (see **Figure 9c and 9d**). Note, however, that several other strong accelerations (during locomotion and around turn off) are not systematically followed by a strong increase in firing at a similar lag. We next plotted the average acceleration of the animals as they traversed the T-maze. Here, we found no consistent direction of correlation regardless of lag (**Figure 9e and 9f**).

Thus, acceleration can account for up to 5% of the variability in the activity of the most highly correlated cells, and on average across all cells in this data set accounts for less than 1%. No cell had a correlation with acceleration high enough to account for 5% of the variability in the sample.

Three maze comparison

Lastly we compared the average firing rate and average speed in our three different tasks; the standard T-maze, the early tone T-maze, and the non associative T-maze (**Figure 10a**). In the standard T-maze, rats learned to turn left or right in response to a tone that turned on mid run (Barnes, Kubota et al. 2005). In the early tone task (the

data in which the rest of the speed analysis have been performed) rats first heard a 500 millisecond tone that predicted which arm of the maze the reward would be in. In the non associative task, tones played like in the standard T-maze but there was no association between the tone and the location of chocolate. The amount of rewarded trials was yoked to the average amount of reward earned by the standard T-maze rats. Behavioral data for all three tasks to stage 10 are plotted in **Figure 10b**.

The average raw firing rates for cells recorded from rats trained in the non associative T-maze is plotted in **Figure 10d**. These rats did not on average show a clear beginning and end activity. However, rats that were trained on the non associative T-maze showed a trend to run slower than rats trained on the standard T-maze (ANOVA p = .066).

Therefore we plotted the raw average firing rates of cells for the fastest third of rats trained in the non associative T-maze (**Figure 10 c,e**). In Figure 9f we directly compared the raw average firing rates of cells recorded in all three tasks during 10 days of overtraining. The mean run time for the 10 days of overtraining plotted for the associative rats was 2.4 s, the non associative rats' mean run time during these stages was 3.0 s. The early tone rats mean run time was 3.1 s. Therefore, the fastest set of rats (the standard T-maze rats) had neither the weakest nor the strongest beginning and end pattern. We also made the comparison using only the fastest third of the rats trained in the non-associative task. Here the average run time for the fastest third of the non associative rats was 2.5 seconds. There is still very little beginning and end activity.

While running these animals, and examining VHS recordings, no obvious difference in motor activity was observed. There was, however, one important

behavioral difference. The majority of non associative animals chose to turn one direction more often than the other. As training progressed 9/11 animals turned their preferred direction every time. This is an important caveat to interpreting these results.

Conclusion

In conclusion these analyses point to a definite contribution of speed and/or position to the firing rate patterns observed and an apparently small contribution of acceleration to these patterns. The analysis in which the speed was held constant (**Figure 6**) and the analysis which excluded the ~24% of single units that showed a correlation with speed (**Figure 7**) suggests that speed was not the dominant factor accounting for the patterns that developed with training.

References

Adams, S., R. P. Kesner, et al. (2001). "Role of the Medial and Lateral Caudate-Putamen in Mediating an Auditory Conditional Response Association." Neurobiology of Learning and Memory **76**(1): 106-116.

Aldridge, J. W. and K. C. Berridge (1998). "Coding of Serial Order by Neostriatal Neurons: A "Natural Action" Approach to Movement Sequence." J. Neurosci. **18**(7): 2777-2787.

Aldridge, W. B., Kent. Rosen, Alyssa. (2004). "Basal ganglia neural mechanisms of natural movement sequences." Cananda Journal of Physiological Pharmacology **82**(8-9): 732-9.

Antonius B. Mulder, E. T., Sidney I. Wiener, (2004). "Neurons in hippocampal afferent zones of rat striatum parse routes into multi-pace segments during maze navigation." European Journal of Neuroscience **19**(7): 1923-1932.

Apicella, P., E. Legallet, et al. (1997). "Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states." Experimental Brain Research **116**(3): 456-466.

Apicella, P., T. Ljungberg, et al. (1991). "Responses to Reward in Monkey Dorsal and Ventral Striatum." Experimental Brain Research **85**(3): 491-500.

Atallah, H. E., D. Lopez-Paniagua, et al. (2007). "Separate neural substrates for skill learning and performance in the ventral and dorsal striatum." Nat Neurosci **10**(1): 126-131.

Bailey, K. R. and R. G. Mair (2006). "The Role of Striatum in Initiation and Execution of Learned Action Sequences in Rats." J. Neurosci. **26**(3): 1016-1025.

Barnes, T. D., Y. Kubota, et al. (2005). "Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories." Nature **437**(7062): 1158-1161.

Barto, A. G. (1995). Models of Information Processing in the Basal Ganglia.

Belin, D., S. Jonkman, et al. (2009). "Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction." Behavioural Brain Research **199**(1): 89-102.

Canales, J. J. (2005). "Stimulant-induced adaptations in neostriatal matrix and striosome systems: Transiting from instrumental responding to habitual behavior in drug addiction." Neurobiology of Learning and Memory **83**(2): 93-103.

Carelli, R. M., M. Wolske, et al. (1997). "Loss of Lever Press-Related Firing of Rat Striatal Forelimb Neurons after Repeated Sessions in a Lever Pressing Task." J. Neurosci. **17**(5): 1804-1814.

Chang, Q. and P. E. Gold (2003). "Switching Memory Systems during Learning: Changes in Patterns of Brain Acetylcholine Release in the Hippocampus and Striatum in Rats." J. Neurosci. **23**(7): 3001-3005.

Corbit, L. H. and P. H. Janak (2007). "Inactivation of the Lateral But Not Medial Dorsal Striatum Eliminates the Excitatory Impact of Pavlovian Stimuli on Instrumental Responding." J. Neurosci. **27**(51): 13977-13981.

DeCoteau, W. E., C. Thorn, et al. (2007). "Oscillations of Local Field Potentials in the Rat Dorsal Striatum During Spontaneous and Instructed Behaviors." J Neurophysiol **97**(5): 3800-3805.

Divac, I., H. E. Rosvold, et al. (1967). "Behavioral effects of selective ablation of the caudate nucleus." Journal of comparative and physiological psychology. **63**(2): 184-90.

Eschenko, O. and S. J. Y. Mizumori (2007). "Memory influences on hippocampal and striatal neural codes: Effects of a shift between task rules." Neurobiology of Learning and Memory **87**(4): 495-509.

Featherstone, R. E. and R. J. McDonald (2005). "Lesions of the dorsolateral striatum impair the acquisition of a simplified stimulus-response dependent conditional discrimination task." Neuroscience **136**(2): 387-395.

Flaherty, A. W. and A. M. Graybiel (1994). "Input-output organization of the sensorimotor striatum in the squirrel monkey." J. Neurosci. **14**(2): 599-610.

Freeman, T. B., F. Cicchetti, et al. (2000). "Transplanted fetal striatum in Huntington's disease: Phenotypic development and lack of pathology." Proceedings of the National Academy of Sciences of the United States of America **97**(25): 13877-13882.

Grahn, J. A., J. A. Parkinson, et al. (2009). "The role of the basal ganglia in learning and memory: Neuropsychological studies." Behavioural Brain Research **199**(1): 53-60.

Graybiel, A. M. (1998). "The Basal Ganglia and Chunking of Action Repertoires." Neurobiology of Learning and Memory **70**(1-2): 119-136.

Graybiel, A. M. (2000). "The basal ganglia." Current Biology **10**(14): R509-R511.

Henry H. Yin, B. J. K. B. W. B. (2004). "Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning." European Journal of Neuroscience **19**(1): 181-189.

Hori, Y., T. Minamimoto, et al. (2009). "Neuronal Encoding of Reward Value and Direction of Actions in the Primate Putamen." J Neurophysiol **102**(6): 3530-3543.

Jog, M., Y. Kubota, et al. (1999). "Building neural representations of habits." Science **286**: 1745-1749.

Jog, M. S., Y. Kubota, et al. (1999). "Building Neural Representations of Habits." Science **286**(5445): 1745-1749.

Kao, M. H., A. J. Doupe, et al. (2005). "Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song." Nature **433**(7026): 638-643.

Kelly RM, S. P. (2004). "Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits." Prog Brain Res **143**: 449-59.

Kish SJ, S. K., Hornykiewicz O. (1988). "Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. Pathophysiologic and clinical implications." N Engl J Med. **318**(14): 876-80.

Kitabatake, Y., T. Hikida, et al. (2003). "Impairment of reward-related learning by cholinergic cell ablation in the striatum." Proceedings of the National Academy of Sciences of the United States of America **100**(13): 7965-7970.

Knowlton, B. J., J. A. Mangels, et al. (1996). "A Neostriatal Habit Learning System in Humans." Science **273**(5280): 1399-1402.

Konorski, J. (1967). Integrative activty of the brain. Chicago, University of Chicago Press.

McDonald, R. J. and N. M. White (1994). "Parallel information processing in the water maze: Evidence for independent memory systems involving dorsal striatum and hippocampus." Behavioral and Neural Biology **61**(3): 260-270.

McGeorge, A. J. and R. L. M. Faull (1989). "The organization of the projection from the cerebral cortex to the striatum in the rat." Neuroscience **29**(3): 503-537.

Mink, J. W. (1996). "THE BASAL GANGLIA: FOCUSED SELECTION AND INHIBITION OF COMPETING MOTOR PROGRAMS." Progress in Neurobiology **50**(4): 381-425.

Mizumori, S. J., O. Yeshenko, et al. (2004). "Parallel processing across neural systems: implications for a multiple memory system hypothesis." Neurobiol Learn Mem **82**(3): 278-98.

Murer, M. G., K. Y. Tseng, et al. (2002). "Brain Oscillations, Medium Spiny Neurons, and Dopamine." Cellular and Molecular Neurobiology **22**(5 - 6): 611-632.

Neal J. Cohen, H. E. B. S. D. S. C. (1985). "Different Memory Systems Underlying Acquisition of Procedural and Declarative Knowledge^a." Annals of the New York Academy of Sciences **444**(Memory Dysfunctions: An Integration of Animal and Human Research From Preclinical and Clinical Perspectives): 54-71.

Olveczky, B. P., A. S. Andalman, et al. (2005). "Vocal experimentation in the juvenile songbird requires a basal ganglia circuit." PLoS Biol. **3**: e153-e153.

Packard, M. G. and B. J. Knowlton (2002). "Learning and memory functions of the basal ganglia." Annual Review of Neuroscience **25**: 563-593.

Packard, M. G. and J. L. McGaugh (1996). "Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning." Neurobiology of Learning and Memory **65**(1): 65-72.

Reynolds, J. N. J., B. I. Hyland, et al. (2001). "A cellular mechanism of reward-related learning." Nature **413**(6851): 67-70.

Rosenbaum, D. A., S. B. Kenny, et al. (1983). "Hierarchical control of rapid movement sequences." Journal of Experimental Psychology: Human Perception and Performance **9**(1): 86-102.

Saka, E. and A. M. Graybiel (2003). "Pathophysiology of Tourette's syndrome: striatal pathways revisited." Brain and Development **25**(Supplement 1): S15-S19.

Sakai, K., K. Kitaguchi, et al. (2003). "Chunking during human visuomotor sequence learning." Experimental Brain Research **152**(2): 229-242.

Samejima, K., Y. Ueda, et al. (2005). "Representation of Action-Specific Reward Values in the Striatum." Science **310**(5752): 1337-1340.

Sato, K., C. Sumi-Ichinose, et al. (2008). "Differential involvement of striosome and matrix dopamine systems in a transgenic model of dopa-responsive dystonia." Proceedings of the National Academy of Sciences **105**(34): 12551-12556.

Schmitzer-Torbert, N. C. and A. D. Redish (2008). "Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype." Neuroscience **153**(2): 349-360.

Schultz, W. (1998). "Predictive Reward Signal of Dopamine Neurons." J Neurophysiol **80**(1): 1-27.

Schultz, W., P. Dayan, et al. (1997). "A Neural Substrate of Prediction and Reward." Science **275**(5306): 1593-1599.

Schultz, W. and R. Romo (1992). "Role of Primate Basal Ganglia and Frontal-Cortex in the Internal Generation of Movements .1. Preparatory Activity in the Anterior Striatum." Experimental Brain Research **91**(3): 363-384.

Siegert, R. J. T., Kathryn D.; Weatherall, Mark; Abernethy, David A. (2006). "Is implicit sequence learning impaired in Parkinson's disease? A meta-analysis." Neuropsychology **20**(4): 490-495.

Smith, A. C., L. M. Frank, et al. (2004). "Dynamic Analysis of Learning in Behavioral Experiments." J. Neurosci. **24**(2): 447-461.

Tang, C., A. P. Pawlak, et al. (2007). "Changes in activity of the striatum during formation of a motor habit." European Journal of Neuroscience **25**(4): 1212-1227.

Tang, C. C., D. H. Root, et al. (2009). "Decreased Firing of Striatal Neurons Related to Licking during Acquisition and Overtraining of a Licking Task." J. Neurosci. **29**(44): 13952-13961.

Tepper, J. M., T. Koós, et al. (2004). "GABAergic microcircuits in the neostriatum." Trends in Neurosciences **27**(11): 662-669.

Tremblay, L., J. R. Hollerman, et al. (1998). "Modifications of Reward Expectation-Related Neuronal Activity During Learning in Primate Striatum." J Neurophysiol **80**(2): 964-977.

Trytek, E. S., I. M. White, et al. (1996). "Localization of motor- and nonmotor-related neurons within the matrix-striosome organization of rat striatum." Brain Research **707**(2): 221-227.

West, M. O., R. M. Carelli, et al. (1990). "A region in the dorsolateral striatum of the rat exhibiting single-unit correlations with specific locomotor limb movements." J Neurophysiol **64**(4): 1233-1246.

White, N. M. and R. J. McDonald (2002). "Multiple Parallel Memory Systems in the Brain of the Rat." Neurobiology of Learning and Memory **77**(2): 125-184.

Wickens, J. R., J. N. J. Reynolds, et al. (2003). "Neural mechanisms of reward-related motor learning." Current Opinion in Neurobiology **13**(6): 685-690.

Wilson, M. and B. McNaughton (1993). "Dynamics of the hippocampal ensemble code for space." Science **261**(5124): 1055-1058.

Yeshenko, O., A. Guazzelli, et al. (2004). "Context-dependent reorganization of spatial and movement representations by simultaneously recorded hippocampal and striatal neurons during performance of allocentric and egocentric tasks." Behav Neurosci **118**(4): 751-69.

Yin, H. H. and B. J. Knowlton (2006). "The role of the basal ganglia in habit formation." Nat Rev Neurosci **7**(6): 464-476.

Yin, H. H., B. J. Knowlton, et al. (2004). "Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning." Eur. J. Neurosci. **19**: 181-189.

Yin, H. H., S. P. Mulcare, et al. (2009). "Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill." Nat Neurosci **12**(3): 333-341.

Figures

A

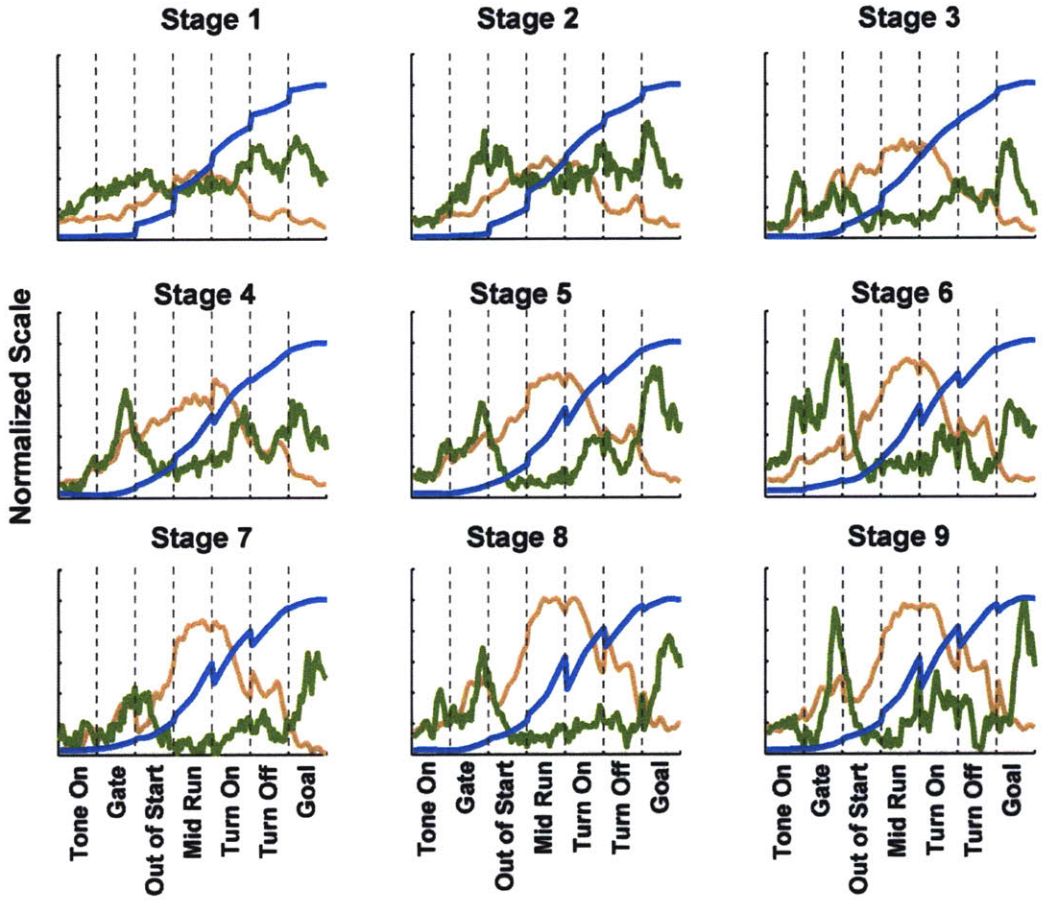


Figure 1. Normalized firing rate (green), normalized speed (orange) and position (blue) as training progresses. The averaged firing rate and speed are plotted in 800 millisecond time epochs around key task events are plotted.

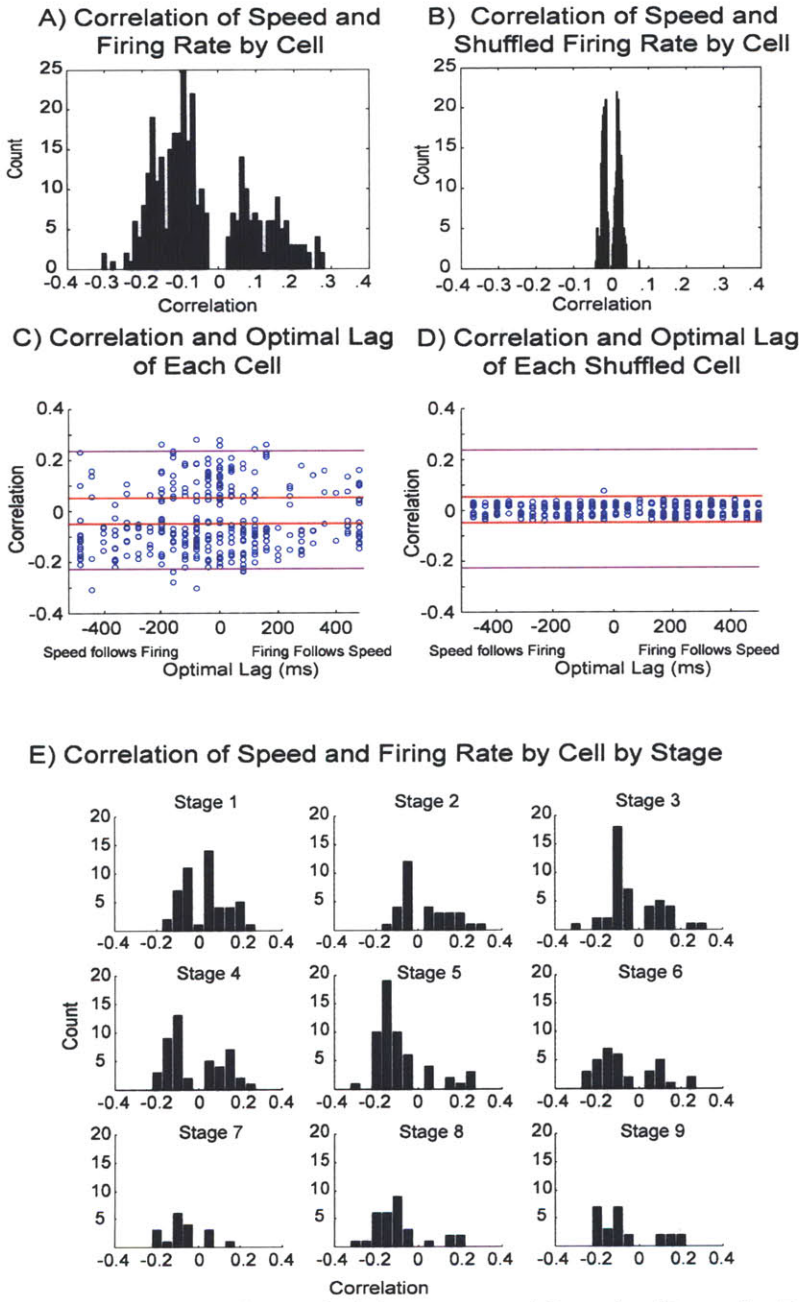


Figure 2. Comparison of Firing Rate and Speed of Rats for Task Responsive Cells. A) Correlation of each task responsive cell's firing rate with speed. For each cell the correlation found at the time lag producing the highest correlation was used. B) Correlation of speed and firing rate for a data set in which the firing rate was shuffled. C) Correlations of each task responsive cell plotted against the optimal lag (in ms). Red lines denote less than p=.01. Pink lines represent r^2 of .05 (5% of variability found in the firing rate can be accounted for by changes in speed). D) The same graph as in C with the shuffled data set. E) Correlation of speed and firing rate for each cell for each stage showing that this relationship is relatively consistent throughout training.

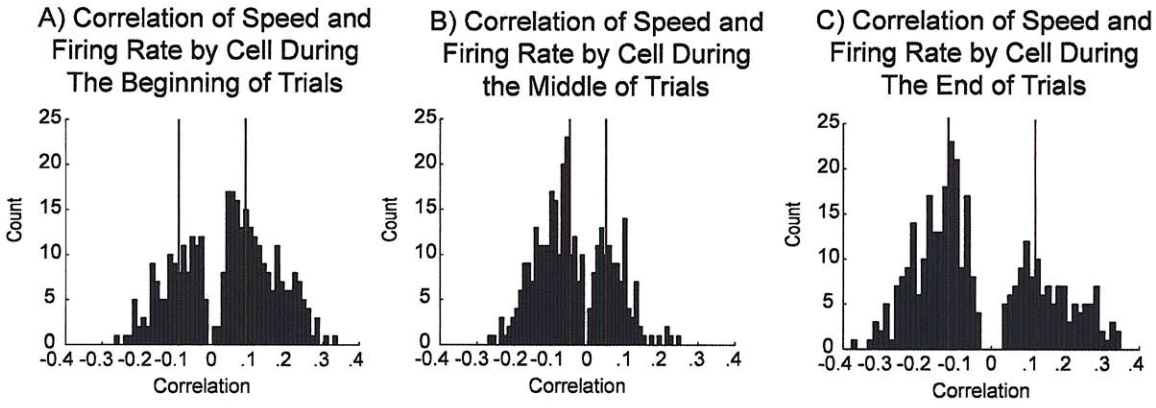


Figure 3. Correlation of Speed and Firing Rate by Cell During the A) Beginning, B) Middle and C) End of Trial. The beginning was defined as 400 ms before tone on to 400 ms after gate was lowered. The middle of the trial was defined from 400 ms after gate was lowered to 400 ms after turn off. The end of the trial was defined from 400 ms after turn off to 400 ms after goal. Red bars indicate less than 1% occurrence by chance. For each cell the correlation found at the time lag (in a -200 millisecond to 200 millisecond window) producing the highest correlation was used.

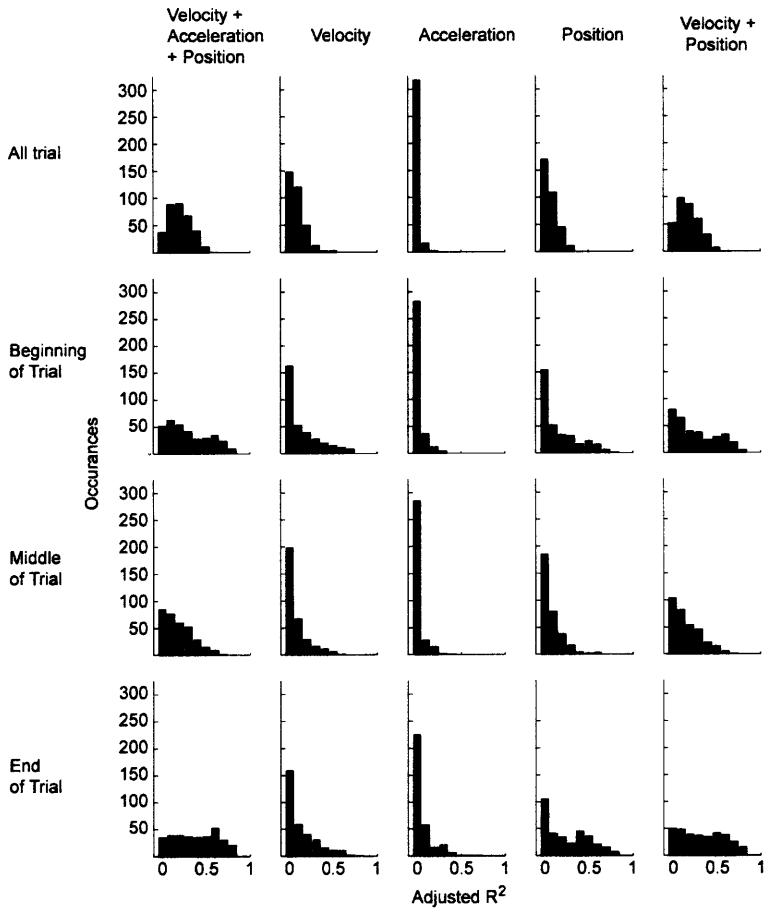


Figure 4. Histograms of adjusted R squared values for regression with firing rate of cells in the dorsal lateral striatum. 336 single putative medium spiny cells from all stages of learning were used. Data were first averaged of all trials in a session.

Table 1. Percent of Variance in Average Firing Rate of Cells in the Striatum That Can Be Explained by Speed, Position, and Acceleration (Adjusted R squared): Single Cell Analysis

	All Together	Speed Alone	Position Alone	Acceleration Alone
All of Task	25%	**12%**	**11%**	2%
Beginning	34%	**18%**	**20%**	5%
Middle	23%	**11%**	**11%**	4%
End	44%	**17%**	**30%**	9%

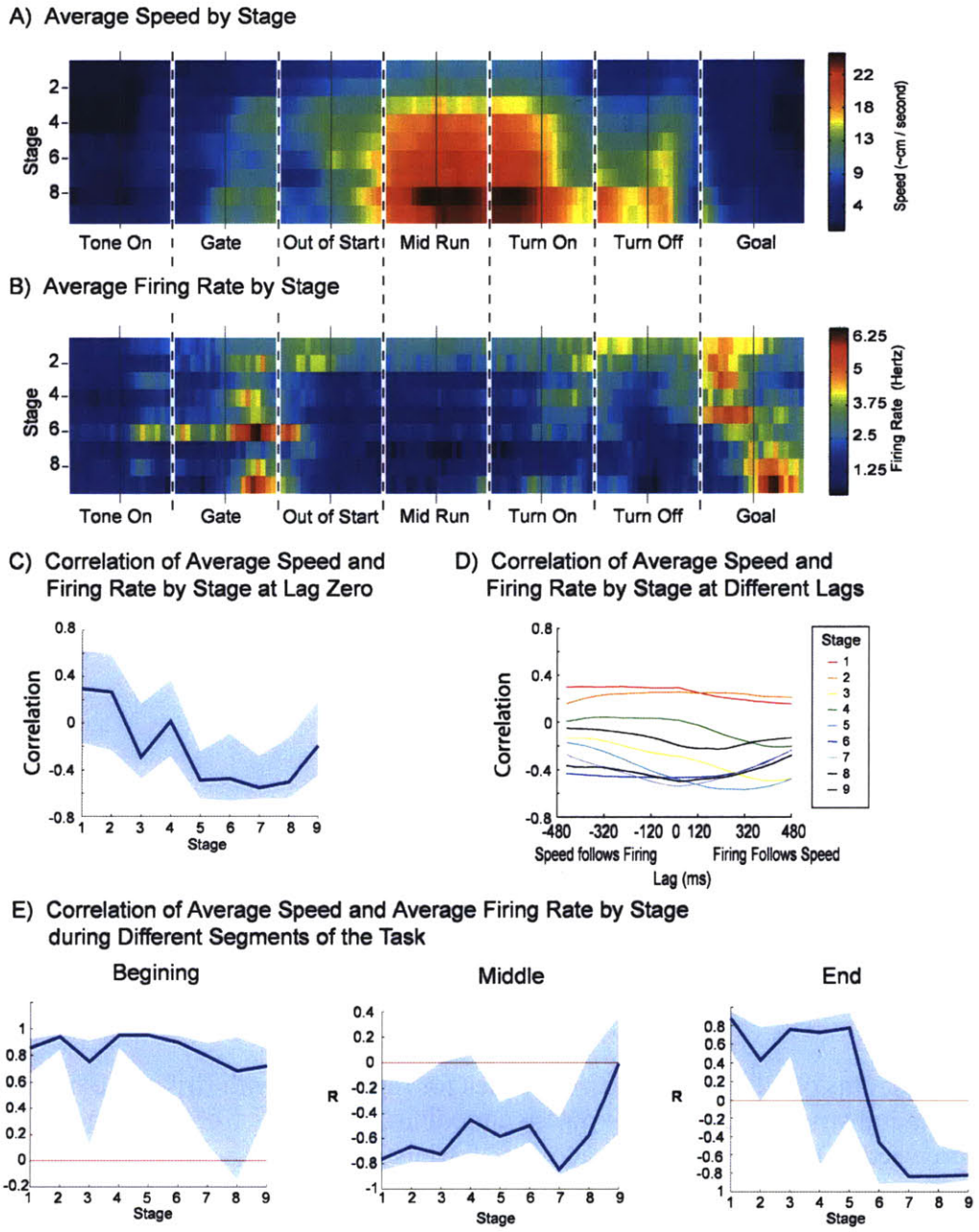


Figure 5. Comparison of average firing rate and average speed of rats. A) Average speed as training progresses. The speed for each rat was calculated and multiplied by the amount of task responsive cells contributed by that animal for that stage of learning. B) Firing rate as training progressed for task responsive cells. Data plotted for 800 milliseconds around each task event in 40 millisecond bins. C) Correlation of average speed and average firing rate by stage. Shaded areas represent 95% confidence intervals. Note that the averaged speed and firing rate do not show a consistent direction of significant correlation. D) Correlation of average speed and average firing rate by stage over different lags .E) Same as in D but broken down into beginning, middle and end.

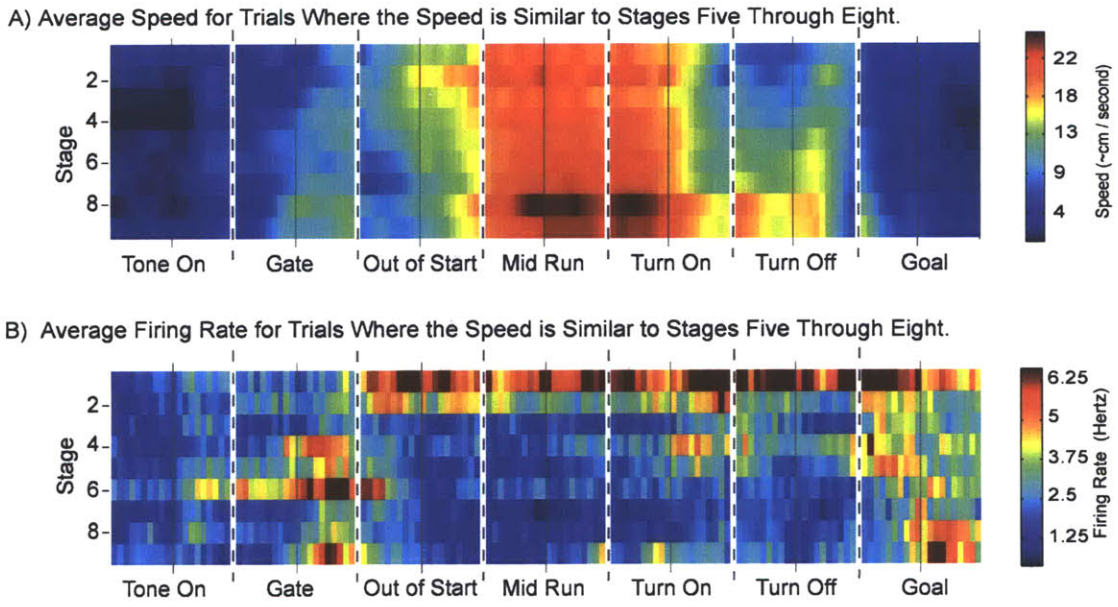


Figure 6. Correlation of firing rate and speed for trials with similar speed. A) Plot of speed for included trials. Note the similarity of the speed traces between stages. For each stage, 246 trials were randomly selected from those trials whose speed for the 800 ms around mid run was at least 1 standard deviation below the mean speed around mid run for stages five through eight. Other time periods were also used to determine if speed was similar and the results were similar. B) Plot of firing rate for included trials demonstrating the changes in firing rate as training progresses.

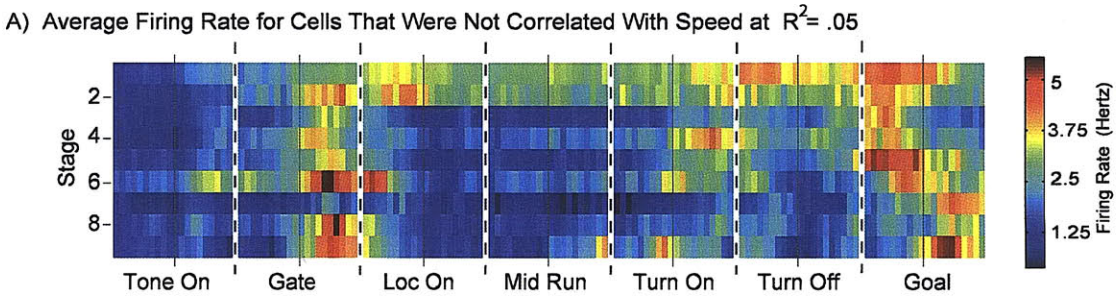


Figure 7. Average firing rate of cells that were not correlated with speed at R^2<.05. 76% of cells remained in the analysis. Data plotted for 800 milliseconds around each task event in 40 millisecond bins.

Table 2. Percent of Variance in Average Firing Rate of Cells in the Striatum That Can Be Explained by Speed, Position, and Acceleration (Adjusted R squared): Collapsed Over Learning Stages-- Ensemble Analysis

	All Together	Speed Alone	Position Alone	Acceleration Alone	Speed + Position	Speed + Position Correlation$
All of Task	36%	**15%**	**21%**	4%	36%	*
Beginning	92%	**88%**	**71 %**	27%	89%	87%
Middle	63%	**40%**	**16%**	27 %	47%	4%
End	32%	*	*	*	*	81%

* indicates not significant to a p=.05 level

$ adjusted R squared of the relationship between speed and position

Table 3. Percent of Variance in Average Firing Rate of Cells in the Striatum That Can Be Explained by Speed, Position, and Acceleration (Adjusted R squared): Average of Nine Stages -- Ensemble Analysis

	All Together	Speed Alone	Position Alone	Acceleration Alone	Speed + Position	Speed + Position Correlation$
All of Task	35%	**14%**	**19%**	*	34%	*
Beginning	74%	**60%**	**59%**	6%	74%	77%
Middle	49%	**32%**	**15%**	11%	40%	8%
End	78%	**43%**	**63%**	14%	73%	69%

* indicates not significant to a p=.05 level

$ adjusted R squared of the relationship between speed and position

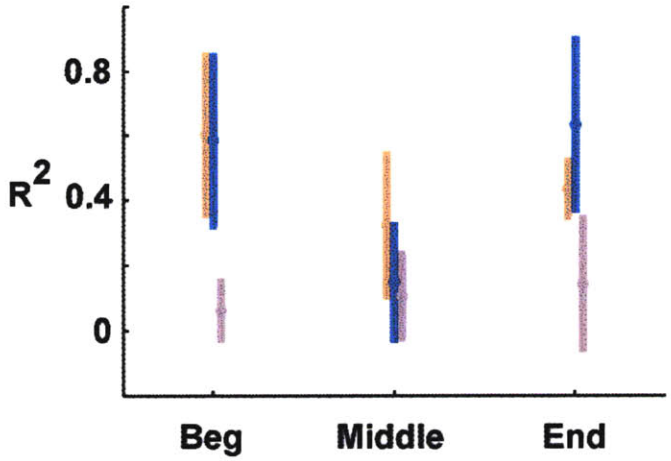


Figure 8. Average adjusted R squared of stages. Average percent of variability for nine stages of learning that can be accounted for in firing rate by speed (orange), acceleration (pink) and position (blue) .Error bars indicate the standard deviation.

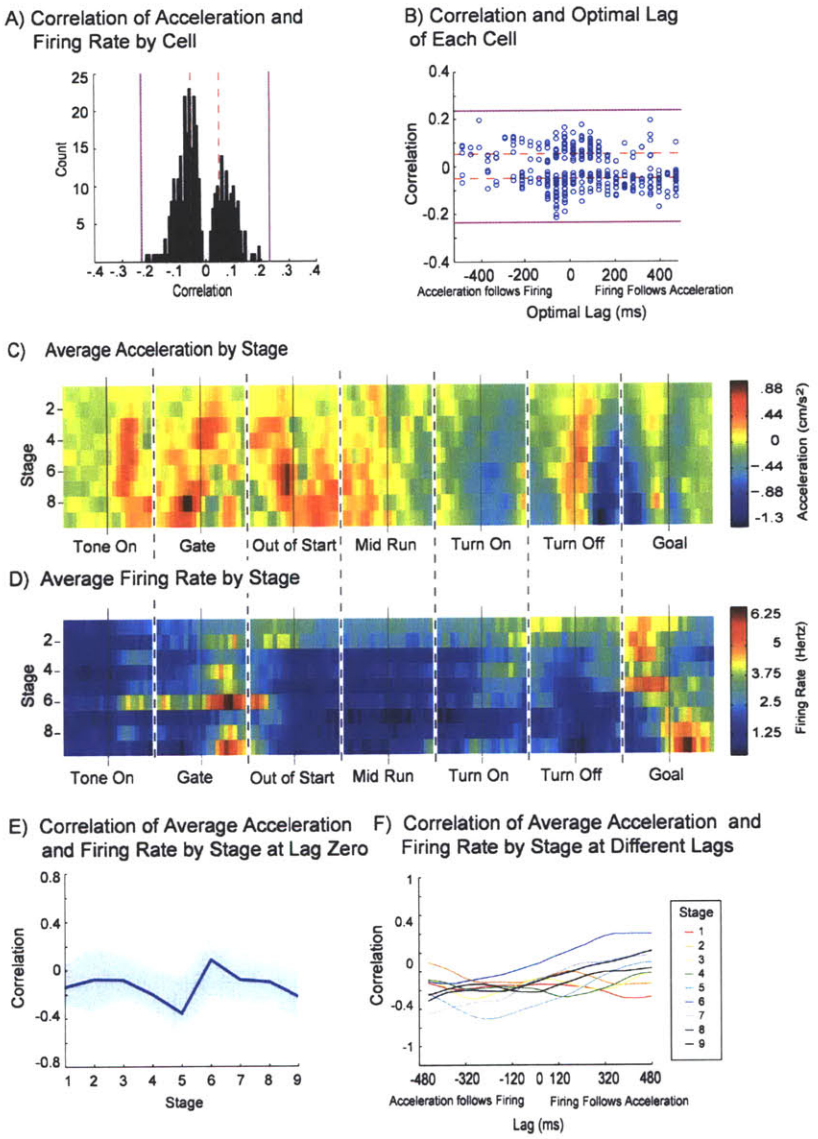


Figure 9. Comparison of firing rate and acceleration of rats for task responsive cells. A) Correlation of each task responsive cell's firing rate with acceleration. For each cell the correlation found at the time lag producing the highest correlation was used. B) Correlations of each task responsive cell plotted against the optimal lag (in ms). Red lines denote less than p=.01. Pink lines represent r^2 of .05 (5% of variability found in the firing rate can be accounted for by changes in acceleration). C) Average acceleration as training progresses. The acceleration for each animal was calculated and multiplied by the amount of task responsive cells contributed by that animal for that stage of learning. D) Firing rate as training progressed for task responsive cells. Data plotted for 800 milliseconds around each task event in 40 millisecond bins. E) Correlation of average acceleration and average firing rate by stage. Shaded areas represent 95% confidence intervals. Note that the averaged acceleration and firing rate do not show a consistent direction of significant correlation. F) Correlation of average acceleration and average firing rate by stage over different time lags.

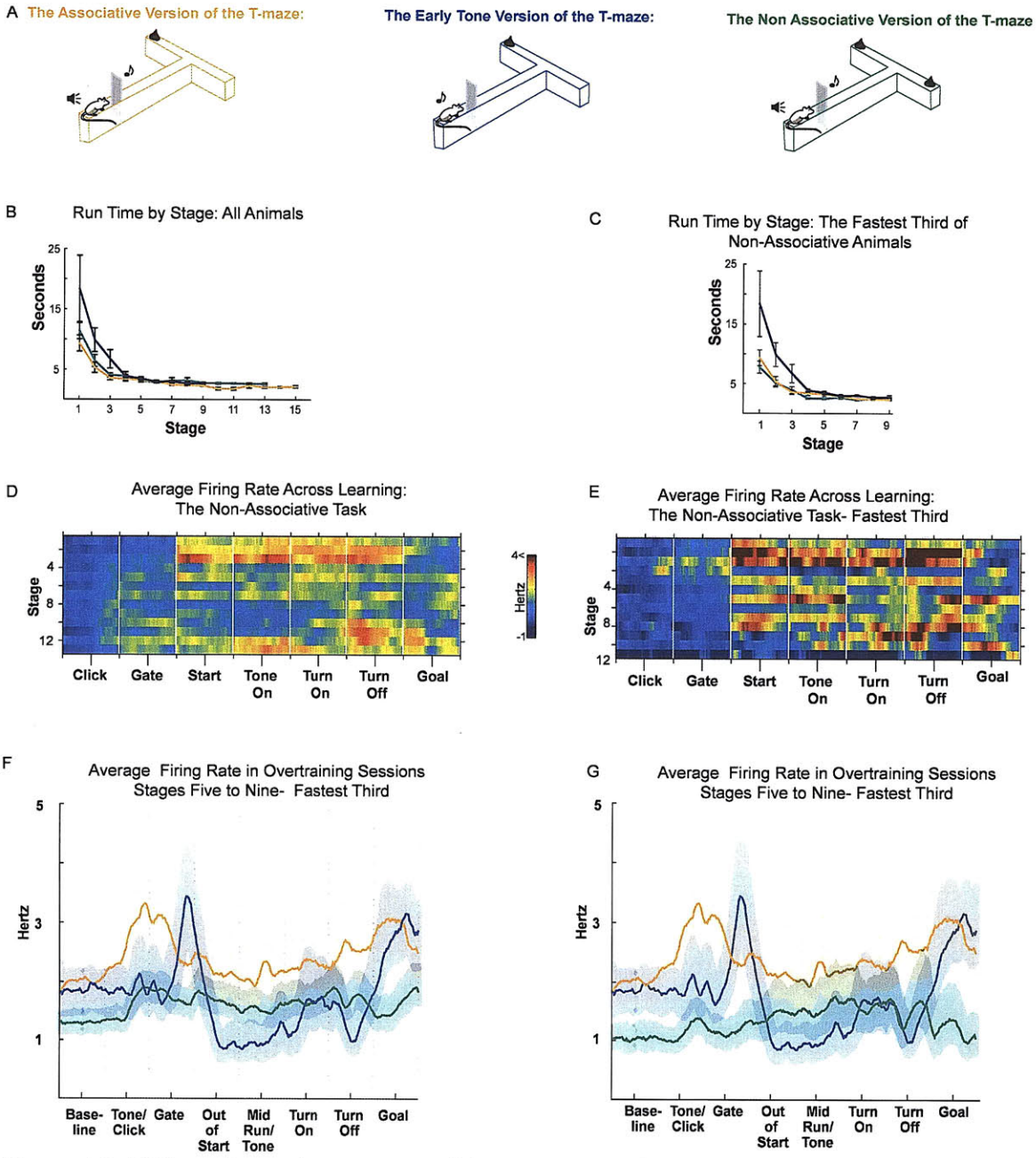


Figure 10. Effect of speed on average firing rate. A) The three different tasks; The associative task (Orange) and the early tone task (Blue) and the non associative task (Green). B) The run times of the three groups of rats. Note that the early tone rats take longer to complete trials, especially on early sessions. C) Run times of all three groups of rats with only the fastest third included for the non associative task. D) The average firing rate across learning for the non associative groups of rats. E) The average firing rate across learning for the fastest third of the non-associative rats. F) The average firing rate during the first 10 days of overtraining, colors as previously described. G) The average firing rate during the first 10 days of overtraining. For the non associative group only the fastest third of rats are included. The mean run time of the non associative rats was, colors as previously described.