Multiple Dopamine Systems: Weal and Woe of Dopamine

Mitsuko Watabe-Uchida and Naoshige Uchida

Center for Brain Science, Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

Correspondence: mitsuko@mcb.harvard.edu; uchida@mcb.harvard.edu

The ability to predict future outcomes increases the fitness of the animal. Decades of research have shown that dopamine neurons broadcast reward prediction error (RPE) signals—the discrepancy between actual and predicted reward—to drive learning to predict future outcomes. Recent studies have begun to show, however, that dopamine neurons are more diverse than previously thought. In this review, we will summarize a series of our studies that have shown unique properties of dopamine neurons projecting to the posterior "tail" of the striatum (TS) in terms of anatomy, activity, and function. Specifically, TS-projecting dopamine neurons are activated by a subset of negative events including threats from a novel object, send prediction errors for external threats, and reinforce avoidance behaviors. These results indicate that there are at least two axes of dopamine-mediated reinforcement learning in the brain—one learning from canonical RPEs and another learning from threat prediction errors. We argue that the existence of multiple learning systems is an adaptive strategy that makes possible each system optimized for its own needs. The compartmental organization in the mammalian striatum resembles that of a dopamine-recipient area in insects (mushroom body), pointing to a principle of dopamine function conserved across phyla.

Since what seems to be the same object may be now a genuine food and now a bait; since in gregarious species each individual may prove to be either the friend or the rival, according to the circumstances, of another; since any entirely unknown object may be fraught with weal or woe, Nature implants contrary impulses to act on many classes of things, and leaves it to slight alterations in the conditions of the individual case to decide which impulse shall carry the day. Thus, greediness and suspicion, curiosity and timidity, coyness and desire, bashfulness and vanity, sociability and pugnacity, seem to shoot over into each other as quickly, and to remain in as unstable equilibrium, in the higher birds and mammals as in man. They are all impulses, congenital, blind at first, and productive of motor reactions of a rigorously determinate sort.

William James, *The Principles of Psychology*, 1890

In natural environments, animals have multiple needs for survival—eating, drinking, avoiding dangers such as predators, and socializing. The above quote describes battles between various competing impulses (James 1890). In selecting an action, animals must balance different needs. The ability to choose an appropriate action in these situations depends critically on the animal's ability to predict consequences of taking particular actions. Animals have various innate mechanisms to detect potential rewards or threats. In ever-changing environments, however, it is the ability to learn from experiences that enhances fitness of the animal.

It has been thought that the neurotransmitter dopamine plays a critical role in learning to predict future outcomes. Although largely confined to a few small midbrain nuclei such as the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc), dopamine-producing neurons project diffusely throughout the brain. Because of this unique anatomical feature, dopamine neurons are well-positioned to broadcast a specific signal to the rest of the brain. Combined with earlier studies pointing to the role of dopamine in reward (Olds and Milner 1954; Wise 2004), studies of dopamine neurons have provided crucial insights into global algorithms by which the brain learns from reward. In this review, we will first describe previous studies on "canonical" dopamine involved in reward-based learning and then novel results pointing to the diversity of dopamine neurons or the idea of multiple dopamine systems.

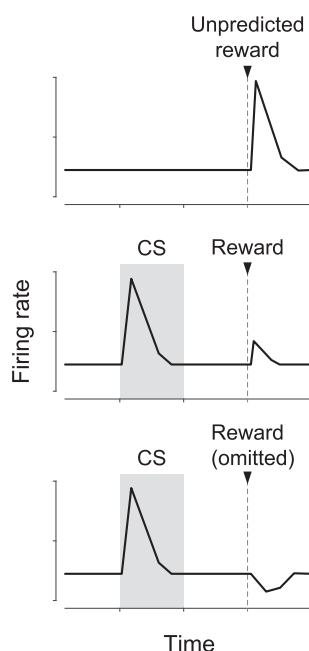## CANONICAL DOPAMINE SIGNALS: REWARD PREDICTION ERRORS

In the 1970s, psychological studies of animal learning indicated that associative learning is driven by prediction errors—the discrepancy between actual and predicted outcome (Kamin 1969; Rescorla and Wagner 1972). When the actual outcome is different from the predicted one, the prediction should be updated. When the prediction is accurate (i.e., when there is no prediction error), no learning will occur. Researchers in machine learning found that prediction error–based learning provides an efficient algorithm in computers that learn from trial and error (Sutton and Barto 1998). In early 1990s, one of such algorithms—temporal difference (TD) learning—achieved human-level performance in a complex board game (backgammon) (Tesauro 1995). More recently, a variant of TD learning algorithms (Q-learning; Watkins and Dayan 1992), combined with deep learning, has achieved human-level performance in far more complex games (Mnih et al. 2015).

A breakthrough in neuroscience came when Wolfram Schultz and colleagues recorded the activity of putative

dopamine neurons in monkeys. It was noticed that the activity of dopamine neurons has remarkable resemblance to the type of RPEs used in TD learning algorithms (Schultz et al. 1997; Bayer and Glimcher 2005). In classical conditioning paradigms in which a cue predicts reward, TD RPEs are distinguished by the following three features (Fig. 1):

1. *Activation by reward-predictive cues.* Dopamine neurons are activated by cues that reliably predict reward. The magnitude of cue-evoked response scales with expected values of future reward.

2. *Expectation-dependent reduction of reward response.* Dopamine neurons are activated by unpredicted reward. Their response to reward is, however, reduced when the reward is predicted by a preceding cue.

3. *Reward omission "dip."* When predicted reward is omitted, dopamine neurons reduce their activity below their baseline firing.

Since then, these firing patterns have been observed in a variety of species and experimental conditions (Oleson et al. 2012; Schultz 2013; Watabe-Uchida et al. 2017). However, as discussed below, whether all dopamine neurons convey TD RPE-like signals remained hotly debated. One difficulty in these studies had been that dopamine neurons were identified using indirect methods based on spike waveform and baseline firing rate (Ungless and Grace 2012). To unambiguously identify dopamine neurons during recording, we tagged dopamine neurons with light-gated cation channel, channelrhodopsin-2 (Boyden et al. 2005; Lima et al. 2009), and identified them based on their responses to light (Cohen et al. 2012). Using this method, we have characterized the activity of dopamine neurons in the lateral VTA in classical conditioning paradigms in mice. Our data showed that optogenetically identified dopamine neurons in the VTA show very similar response properties among each other, largely consistent with TD RPEs (Cohen et al. 2012, 2015; Tian and Uchida 2015; Matsumoto et al. 2016; Starkweather et al. 2017, 2018). In one line of work, we found that their responses to reward were reduced by reward expectation in a purely subtractive fashion, and each neuron's response functions, both for unexpected and expected reward, were scaled versions of one another (Eshel et al. 2015, 2016), demonstrating a remarkable homogeneity in dopamine signals originating in this region. These studies have indicated that the brain employs a prediction error-based learning algorithm, akin to those developed in machine learning. Reinforcement learning theories provide normative perspectives on animal learning and dopamine functions.

NONCANONICAL DOPAMINE RESPONSES

Despite the success of the TD RPE account of dopamine signals, some studies have challenged this "canonical" view. First, some studies have found that at least some dopamine neurons are activated by aversive stimuli in addition to rewarding stimuli (Matsumoto and Hikosaka 2009). This led to the proposal that these dopamine neurons, mainly located in the lateral part of SNc, signal "motivational salience" (the unsigned absolute value of an outcome) and facilitate a behavioral reaction when an important stimulus is detected (Matsumoto and Hikosaka 2009; Bromberg-Martin et al. 2010b). Another deviation from TD RPEs is that some dopamine neurons are activated by novelty (Steinfels et al. 1983; Ljungberg et al. 1992; Horvitz et al. 1997; Rebec et al. 1997; Lak et al. 2016). There have been attempts to incorporate these novelty signals into the reinforcement learning framework: it was proposed that these signals represent a "bonus" to the RPE signals because novelty may be rewarding itself ("novelty bonus") or signal potential reward ("shaping bonus") (Kakade and Dayan 2002). However, functions of these noncanonical dopamine signals have not been shown experimentally, and, therefore, whether motivational salience or novelty bonus well describes functions of dopamine neurons remains unclear.

DIVERSITY—CONNECTIVITY

In addition to showing diversity in their activity patterns (as described above), dopamine neurons differ in terms of gene expression and intrinsic neurophysiological properties (Lacey et al. 1989; Grimm et al. 2004; Lammel et al. 2008; Roeper 2013; Poulin et al. 2014; Lerner et al. 2015). Importantly, these differences tend to correlate with where their axons project to (i.e., their projection targets).

As a foray into the diversity of dopamine neurons, we sought to compare anatomical properties of different populations of dopamine neurons (Watabe-Uchida et al. 2012). We reasoned that because the pattern of activity is largely shaped by their inputs, studying the sources of monosynaptic input may provide not only insights into the


Figure 1. Canonical reward prediction error signals.
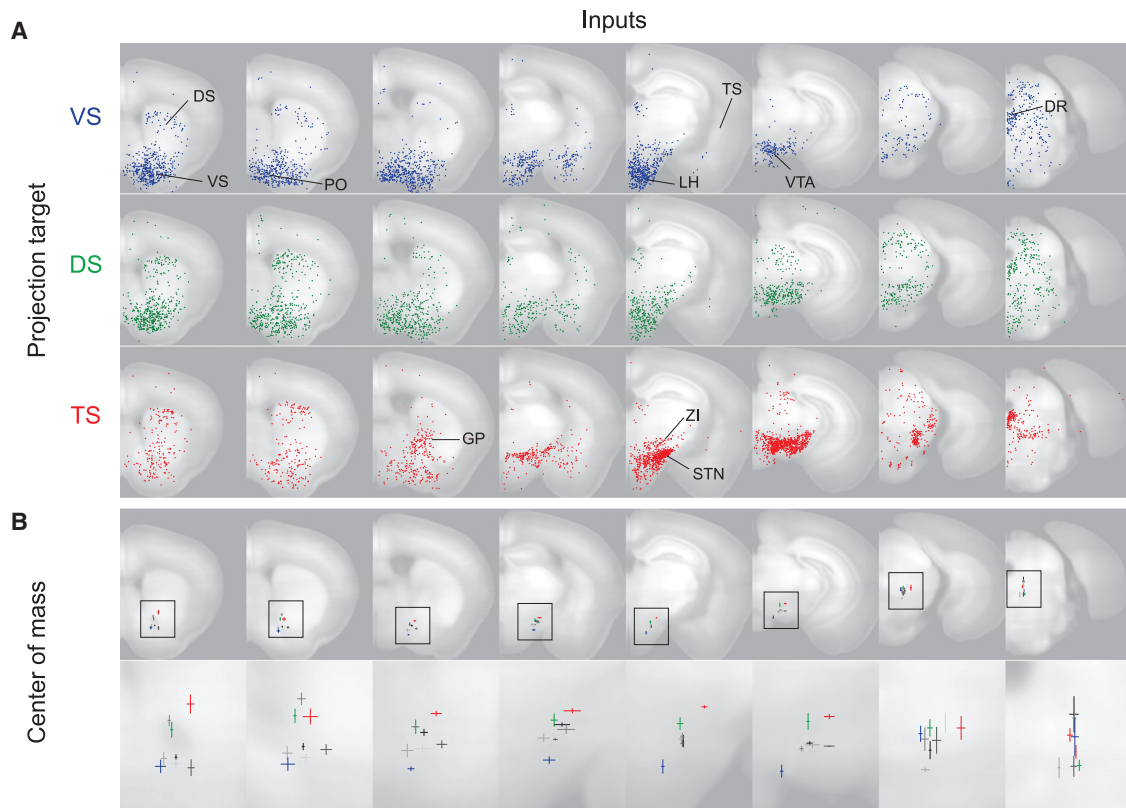
Figure 2. Distribution of monosynaptic inputs to projection-specific dopamine neurons. (*A*) Coronal sections. Monosynaptic inputs to dopamine neurons projecting to the three regions of the striatum are labeled using trans-synaptic rabies virus. (VS) Ventral striatum, (DS) dorsal striatum, (TS) tail of striatum, (VTA) ventral tegmental area, (PO) preoptic area, (LH) lateral hypothalamus, (DR) dorsal raphe, (GP) globus pallidus, (ZI) zona incerta, (STN) subthalamic nucleus. (*Left*) anterior, (*right*) posterior. (*B*) Center of mass of monosynaptic inputs to eight different populations of dopamine neurons. Blue, green, and red correspond to monosynaptic inputs to VS-, DS-, and TS-projecting dopamine neurons. Mean ± SEM. (Portion adapted from Menegas et al. 2015.)

basic mechanism of how dopamine responses are generated but also clues as to the diversity of dopamine neurons. To this goal, we applied a trans-synaptic rabies tracing system (Wickersham et al. 2007) to compare monosynaptic inputs for dopamine neurons in VTA versus SNc (Watabe-Uchida et al. 2012). We found that these dopamine neuron populations receive input from overlapping but distinct sets of brain regions. Although this study distinguished dopamine subpopulations based on their locations of cell bodies, dopamine neurons projecting to different targets are intermingled in these areas. Recent studies have indicated the importance of distinguishing dopamine neurons based on their targets (Roeper 2013). More recent studies, therefore, identified inputs to dopamine neurons separated by their projection targets (Beier et al. 2015; Lerner et al. 2015; Menegas et al. 2015). In our study (Menegas et al. 2015), we combined rabies virus–based tracing with a brain-clearing method (CLARITY) (Chung et al. 2013), light-sheet microscopy (Keller et al. 2010), and automated analysis software. This allowed us to examine dopamine populations projecting to eight different targets. Based on this data set, we found that dopamine neurons projecting to the posterior "tail" of the striatum (TS) have a unique set of inputs compared to other populations projecting to the ventral striatum (VS),

dorsal striatum (DS), globus pallidus, orbitofrontal cortex, medial prefrontal cortex, amygdala, and habenula (Figs. 2 and 3; Menegas et al. 2015). Although the VS was a major source of input to all of the other seven populations, TS-projecting dopamine neurons received little from the VS. Instead, TS-projecting dopamine neurons received relatively larger numbers of input from dorsolaterally shifted regions such as the subthalamic nucleus, zona incerta, and globus pallidus (Fig. 2). These results raised the possibility that TS-projecting dopamine neurons are particularly unique among dopamine neuron populations, and may show different activity patterns as well as functions.

DIVERSITY—ACTIVITY

The aforementioned results indicate the importance of distinguishing dopamine neurons according to their projection targets. Our previous electrophysiological recording using optogenetic identification (Cohen et al. 2012; Eshel et al. 2015, 2016), however, did not allow us to identify their projection sites and mainly targeted VTA but not SNc. Dopamine signals in specific projection sites have been characterized using an electrochemical method (cyclic voltammetry) or a direct measurement of dopamine
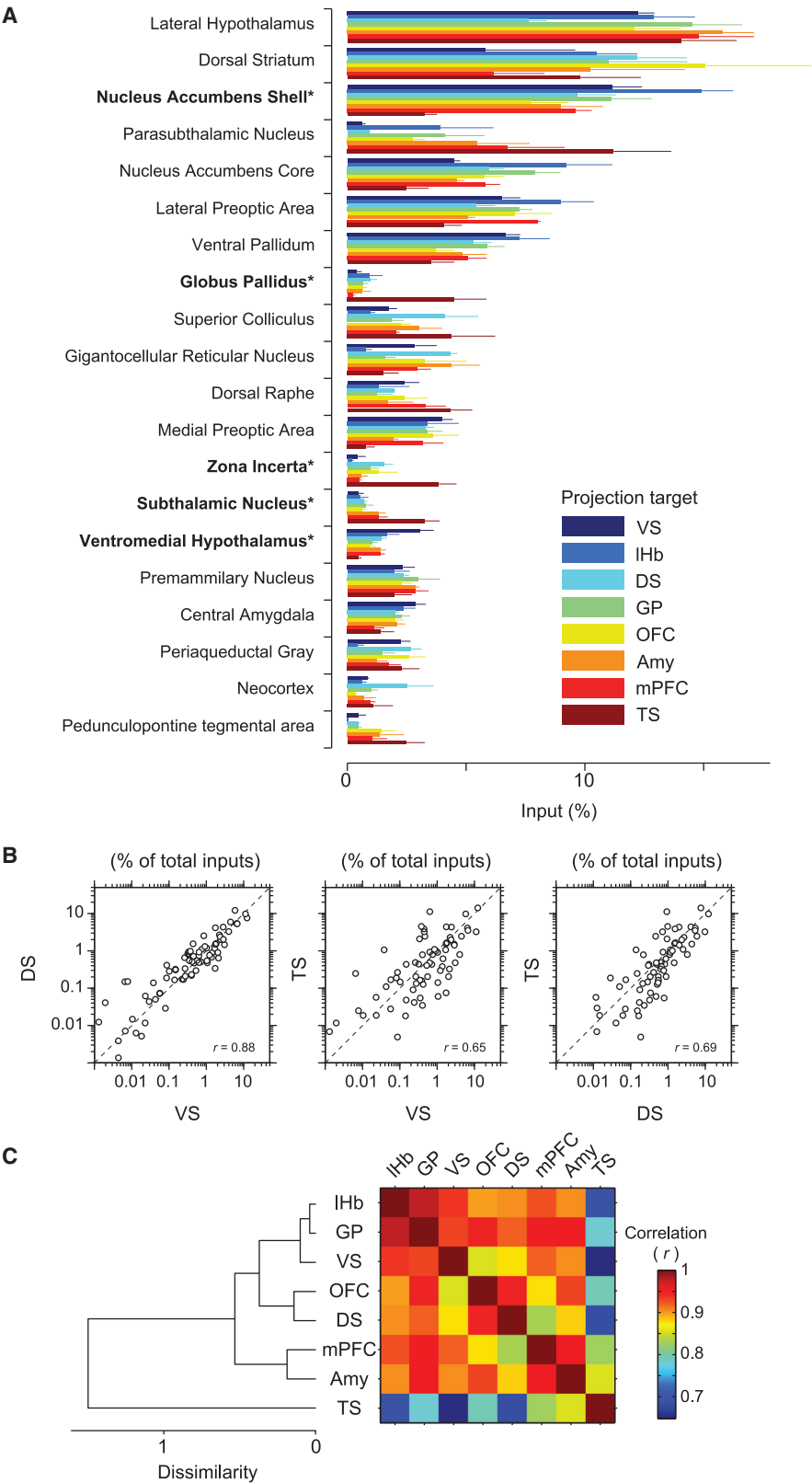
Figure 3. Comparison of the patterns of monosynaptic inputs to projection-specific dopamine neurons. (*A*) Percentage of inputs originating from each area. Top 20 areas are shown. (VS) Ventral striatum, (lHB) lateral habenula, (DS) dorsal striatum, (GP) globus pallidum, (OFC) orbitofrontal cortex, (Amy) amygdala, (mPFC) medial prefrontal cortex, (TS) tail of striatum. (*B*) Correlation analysis between pairs of dopamine neurons populations. (*C*) Correlation matrix between eight dopamine neuron populations. TS-projecting dopamine neurons are an outlier among the eight populations. (Adapted from Menegas et al. 2015.)

using microdialysis. Microdialysis is superior in chemical specificity but slow (on the order of tens of seconds to minutes). Cyclic voltammetry has a faster temporal resolution (tens to hundreds of milliseconds) but it is often difficult to isolate dopamine from other chemicals such as noradrenaline, restricting its application to specific brain areas where noradrenaline is scarce (e.g., VS). Nonetheless, early studies using cyclic voltammetry have provided critical information as to dopamine dynamics in specific targets. Roitman and colleagues measured dopamine concentrations in the VS (nucleus accumbens), and found that reward increased dopamine release, whereas aversive bitter taste decreased it (Roitman et al. 2008), consistent with canonical dopamine responses. Hart and colleagues have provided evidence that the dopamine concentration in the VS faithfully encodes RPEs (Hart et al. 2014). Contrary to these results in the VS, dopamine in other regions (e.g., dorsal striatum) did not necessarily follow RPEs and remained to be clarified (Brown et al. 2011).

More recently, calcium sensor–based methods have been used to monitor dopamine neuron activities in a projection-specific manner. Fiber fluorometry (also called "fiber photometry") is used to measure fluorescent signals through fiber optics (Kudo et al. 1992). Fiber fluorometry, combined with sensitive Ca^{2+} indicators expressed in a cell type–specific manner, now allows one to monitor dopamine neuron population activities at cell bodies as well as at axon terminals (Gunaydin et al. 2014; Lerner et al. 2015; Howe and Dombeck 2016; Menegas et al. 2017). Furthermore, two-photon Ca^{2+} imaging has allowed one to monitor the activity of dopamine axons (Howe and Dombeck 2016) and at cell bodies (Engelhard et al. 2018). Recent studies using these methods have begun to reveal dopamine signals in a projection- and cell type–specific manner (Howe and Dombeck 2016; Kim et al. 2016; Parker et al. 2016; Matias et al. 2017; Menegas et al. 2017, 2018).

In our recent studies (Menegas et al. 2017, 2018), we monitored the activity of midbrain dopamine neurons projecting to the striatum, the major dopamine-recipient area in the brain (Fig. 4). We compared dopamine axon Ca^{2+} signals in four different areas of the striatum—ventral striatum (VS), dorsomedial striatum (DMS), dorsolateral striatum (DLS), and posterior "tail" of the striatum (TS) (Menegas et al. 2017). A genetically encoded Ca^{2+} indicator, GCaMP6, was expressed in dopamine neurons, and calcium signals from axons were collected from fiber optics implanted into the striatal regions in head-fixed mice performing in a classical conditioning paradigm (Fig. 4A).

In the VS, we observed all the three features of RPE-related activities: activation by reward-predictive cues, reduction of reward responses by reward expectation, and a dip in activity caused by omission of predicted reward (Fig. 4B, left). These RPE-related signals reflect outcome values: (1) The response to reward scaled with increasing amounts of water, (2) all negative outcomes that we tested (air puff, bitter taste, and omission of reward) inhibited them, and (3) neutral stimuli (e.g., pure tones with varying intensity) did not evoke notable responses. The reward responses were widespread across the dorsal striatum (DLS and DMS), whereas air puff responses were much weaker, but sometimes positive, in the dorsal striatum (also see Lerner et al. 2015).

In the TS, in stark contrast to the VS (and DMS and DLS), rewards or reward-predictive cues caused little activation of dopamine axons, and varying amounts of water did not modulate the level of activation. It is of note that, in our earlier study (Menegas et al. 2017), we observed significant excitation during water delivery, but our later study found that these responses are diminished in the presence of sounds that masked the noise of water delivery (Menegas et al. 2018). These results together indicated that TS dopamine does not signal reward values. Instead, dopamine axons in TS were strongly activated by air puff (Fig. 4B, right) or loud sound. The level of activation was modulated by the intensity of air puff or sound. Interestingly, TS dopamine axons were not activated by all negative events: They did not respond to bitter taste or reward omission (Fig. 4C). We also found that these unique response properties are present not only at their axons but also at their cell bodies, indicating that these unique responses are not due to local modulations at the axons but reflect cellular activities (Menegas et al. 2018).

Another striking difference between dopamine signals in VS and TS was found during novel odor learning (Fig. 4D; Menegas et al. 2017). At the beginning of new cue–reward associations, VS dopamine axons responded strongly to reward but not to the cue. As the learning proceeds, the magnitude of reward responses gradually decreased while that of cue response increased (Ljungberg et al. 1992; Mirenowicz and Schultz 1994; Stuber et al. 2008; Flagel et al. 2011). These learning-dependent changes occurred over the course of tens of trials in well trained animals (in some cases, even in 1–2 trials; Bromberg-Martin et al. 2010a; Babayan et al. 2018) although when the animal was first trained in an odor-reward association task, these changes occurred in a much longer timescale (over the course of several days) (Menegas et al. 2017). In stark contrast, TS dopamine axons were activated strongly by a novel cue from the very first exposure of a naive animal, even before the animal experienced the associated outcome in the context (Menegas et al. 2017). These response patterns gradually decreased over tens of trials. These responses were observed across different sensory modalities—olfactory, visual, and auditory (Menegas et al. 2018), suggesting that they represent the novelty of sensory stimuli.

DIVERSITY—FUNCTIONS

The above studies revealed unique response properties of TS-projecting dopamine neurons: They are activated by a subset of negative events (e.g., air puff, loud sound) and novel stimuli. What are the functions of TS-projecting dopamine neurons? What do these negative events and novel stimuli have in common? We next addressed these questions (Menegas et al. 2018).

Activation of canonical dopamine neurons has a rewarding effect—it increases the frequency of actions that lead to their activation (Tsai et al. 2009; Witten et al. 2011; Stein-
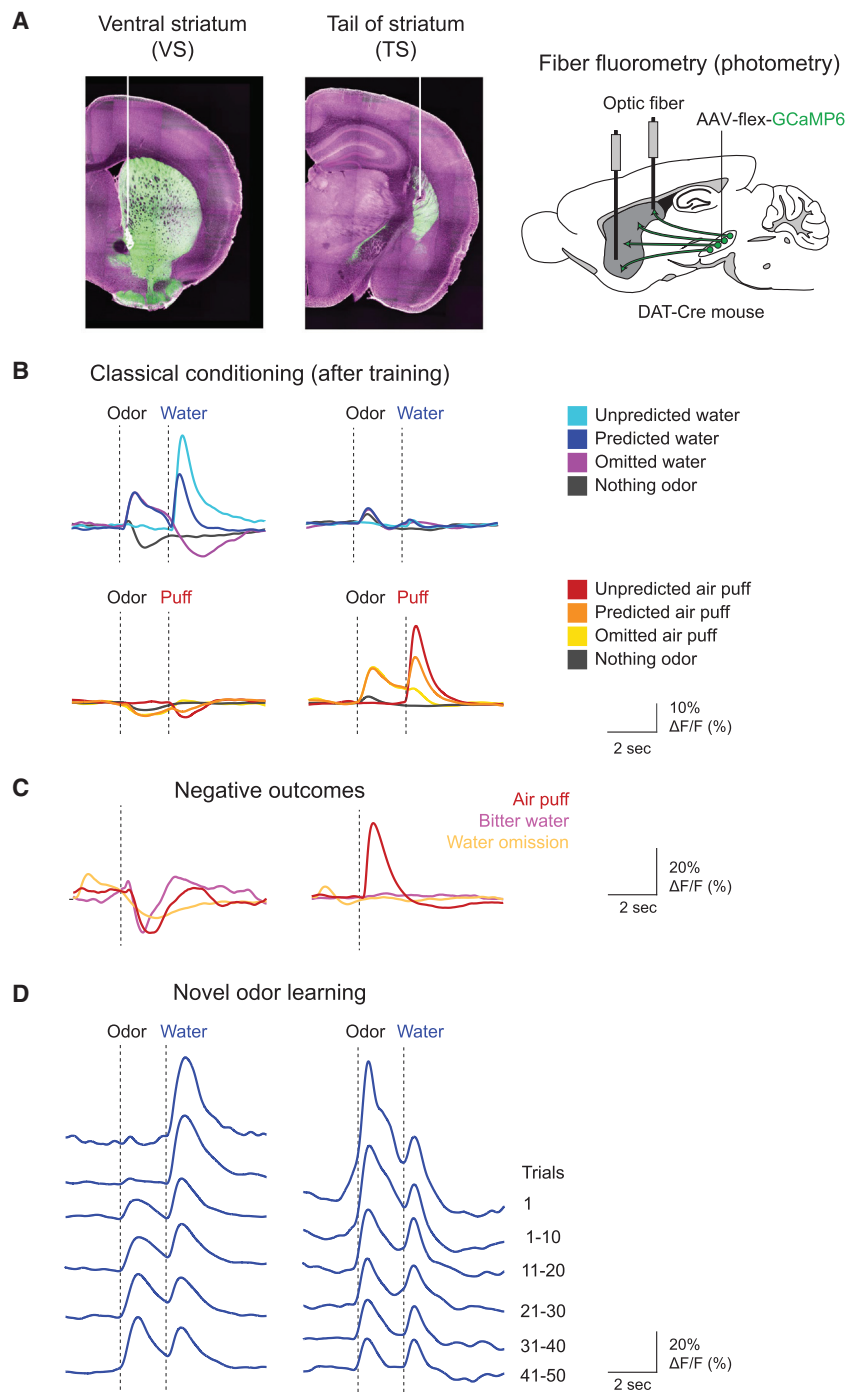
Figure 4. Response properties of dopamine neurons projecting to the ventral striatum (VS) and the tail of striatum (TS). (*A*) Schematic of activity measurement using fiber fluorometry. (*B*) Response patterns in a classical conditioning paradigm. (*Top*) Reward trials with masking sound (Menegas et al. 2018). (*Bottom*) Air puff trials (Menegas et al. 2017). (*C*) Responses to three different negative outcomes (Menegas et al. 2018). (*D*) Changes of activity during novel odor–reward associations (Menegas et al. 2017). (Adapted from Menegas et al. 2017, 2018.)

berg and Janak 2013). Our finding that TS-projecting dopamine neurons are activated by a subset of negative events but not by reward indicated that they have different functions. We therefore first tested the effect of activating TS-projecting dopamine neurons (Fig. 5; Menegas et al. 2018). In a choice assay, mice were trained to nose poke

into the central port and then move to the left or right choice port to obtain reward. In a block of trials, while receiving water reward, mice received a brief optogenetic stimulation of dopamine neurons in VS or TS. Stimulation in the VS biased the animal's choice toward the stimulated side, consistent with a rewarding effect. To our surprise, stimulation
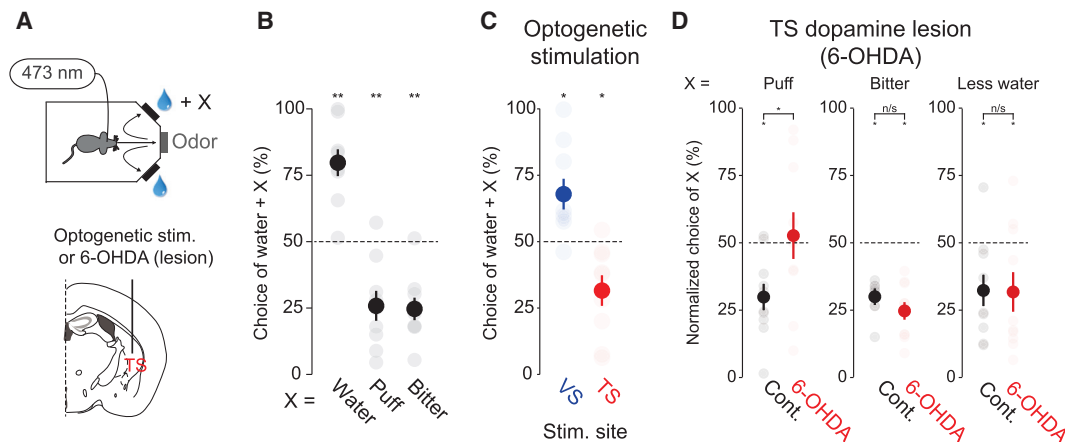
Figure 5. Activation of TS-projecting dopamine neurons reinforce avoidance behavior in a choice task. (*A*) Schematic of experiments. (*B*) Choice bias caused by extra water (Water), air puff (Puff), and bitter taste (Bitter). (*C*) Effects of optogenetic stimulation of VS- or TS-projecting dopamine neurons at the time of water receipt. (*D*) Effects of lesioning TS-projecting dopamine neurons. (Adapted from Menegas et al. 2018.)

of dopamine axons in TS biased the animal's choice away from the stimulated side. Next we performed lesions of TS-projecting dopamine neurons using a neurotoxin, 6-hydroxydopamine. In the choice assay, if one side is associated with negative outcomes such as air puff, bitter taste, or a reduction of water, normal animals bias their choices away from it. After lesioning TS-projecting dopamine neurons, however, mice lost their bias away from air puff, although their baseline activities remained unimpaired. Interestingly, the lesioned mice were able to learn to avoid the side associated with bitter taste or a reduction of reward, mirroring the response properties of TS-projecting dopamine neurons. Furthermore, microinjection of a dopamine D1 receptor antagonist into TS impaired their ability to learn from air puff. Of note, in the optogenetic stimulation experiment, the timing of optogenetic stimulation of TS-projecting dopamine neurons (during outcome) are segregated from the timing of choice. Thus, these results are consistent with the possibility that dopamine in TS acts as a reinforcement signal that alters their future behaviors to facilitate avoidance of threatening stimuli, although we cannot completely rule out the possibility that it also modulates ongoing behaviors.

The above results showed a critical role of dopamine in the TS in avoidance learning. Aforementioned results also showed that TS-projecting dopamine neurons are activated by novelty. These novelty responses are difficult to understand in terms of "novelty bonus" or "motivational salience." First, novelty bonus refers to an extra value added to an object to facilitate exploration. However, the response properties of TS-projecting dopamine neurons do not encode reward value in the first place. Furthermore, the lack of responses to reward is inconsistent with the representation of motivational salience, which was originally characterized as a response to stimuli of both positive and negative valence.

So what do these novelty responses of TS dopamine represent? What is the relationship between the novelty response and avoidance learning? When animals encoun-

ter a novel object, they often show bouts of careful approach and quick retreat—much like "weal and woe" described by William James (1890). Such behaviors were also characterized as risk assessment behavior. In our experiments (Menegas et al. 2018), mice showed this approach–retreat behavior for a long period of time often lasting >10 min following a first encounter with a novel object (Fig. 6). After lesioning TS-projecting dopamine neurons, their interactions with a novel object were greatly altered. At the very initial interactions, they showed normal, short bouts of approach–retreat behaviors. Subsequently, however, they quickly lost this cautious behavior and showed much longer interactions with the novel object. The number of approaches was unaltered but the duration of interaction per bout was increased after a few minutes. These results indicated that TS-projecting dopamine neurons play a critical role in "maintaining" careful approach–retreat behaviors. Potentially, TS-projecting dopamine neurons specifically signal a threatening aspect of novelty: They are activated when the animal "recognizes" a novel object and act as a reinforcement signal to facilitate retreat behaviors in future bouts, much like a reinforcement signal for avoidance in the choice assay. Consistent with this view, we found that TS-projecting dopamine neurons are activated at the time when the animal is in proximity to a novel object (at the timing of updating), but not when they start approaching. Furthermore, optogenetic stimulation of TS-projecting dopamine neurons recovered short bouts of approach–retreat behaviors against a familiar object. Importantly, short bouts of interactions lasted even after optogenetic stimulation was turned off. These results indicated that dopamine in TS regulates the animal's retreat behaviors during novelty interactions by functioning as a reinforcement signal to maintain quick retreat from a novel object.

In total, these results indicate that TS dopamine may encode external threats and acts as a reinforcement signal to facilitate avoidance behaviors. Hence, the canonical dopamine system and TS dopamine system may regulate
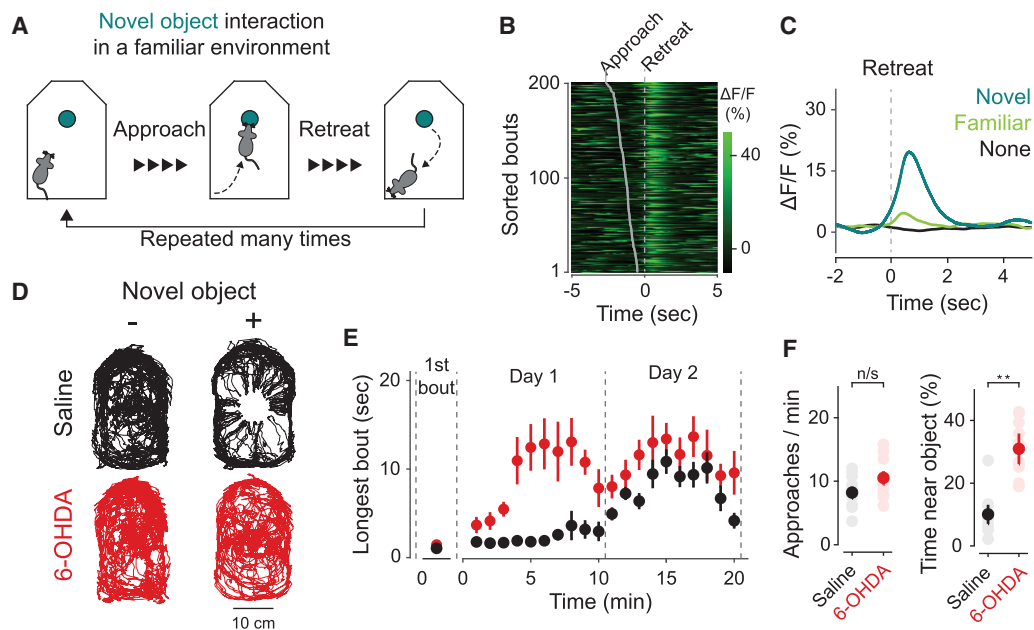
Figure 6. TS-projecting dopamine neurons are involved in maintaining retreat from a novel object. (*A*) Mice show approach–retreat bouts against a novel object. (*B,C*) Activity of TS-projecting dopamine neurons measured using fiber fluorometry. (*D–F*) The effect of lesioning TS-projecting dopamine neurons. (Adapted from Menegas et al. 2018.)

competing drives that lead to reward seeking and threat avoidance, respectively.

ALGORITHM FOR LEARNING FROM DIFFERENT OUTCOMES

As discussed above, VS-projecting dopamine neurons are activated by positive outcomes, whereas they are inhibited by negative outcomes. They show the three features of TD error responses as defined above. During learning, they gradually acquire responses to a novel reward predictor. Activation of VS-projecting dopamine neurons acts as a positive reinforcement. The theoretical framework of reinforcement learning, thus, provides a solid basis for understanding dopamine in VS (see Fig. 8A, left). In contrast, TS dopamine differs from VS dopamine in many ways. First, activation of these neurons facilitated avoidance of events that led to a dopamine release in TS. Second, TS-projecting dopamine neurons are activated by a subset of negative events but not by reward. These responses may represent TD error signals not for reward but for threats; however, they do not strictly follow the predictions of TD learning: They did not show a "dip" when the predicted air puff was omitted and they were strongly activated by novel stimuli before the animal learns whether it is a good predictor of a good or bad outcome. Instead of seeing these "violations" as the evidence against TD learning, these unique response properties can be seen as adaptive changes of TD learning specialized for threat (see Fig. 8A, right), working in concert with the canonical reinforcement learning system in the VS that learns more general outcome values. In natural environments, it is often life-threatening if an animal erroneously approaches a dangerous object (e.g., predator). Thus, it is sensible to increase the "carefulness" or "timid-

ity" against a novel object at the beginning of the interaction (i.e., default of threat prediction is high). Moreover, once the animal learns that a certain stimulus predicts danger, it is sensible not to erase that memory too quickly (i.e., omission of threat does not actively cause weakening of threat prediction). The "violations" of TD error responses by TS dopamine could be seen as achieving conservative and more adaptive approaches for learning about threatening stimuli. Seen in this way, the reinforcement learning framework can still provide a basic framework for understanding TS dopamine: TS dopamine signals threat prediction errors and reinforces threat prediction that in turn maintain avoidance behaviors. In other words, dopamine in VS and TS share common mechanisms at the algorithmic level.

In total, these results indicate that there are at least two axes of reinforcement learning in the striatum, one for learning about outcome values through RPEs (or value prediction errors) and the other for learning about external threats through threat prediction errors. Having separate learning systems would be advantageous because each can have different parameters for learning optimized for different objectives—such as how quickly learning should take place (i.e., learning rate) and how they should react under uncertainty such as when they encounter a novel object.

WHAT IS THE POSTERIOR "TAIL" OF STRIATUM?

Different regions of the striatum receive distinct sets of inputs from the neocortex, thalamus, and amygdala. These inputs are then interfaced with different dopamine signals, and they together constrain functions of each striatal re-
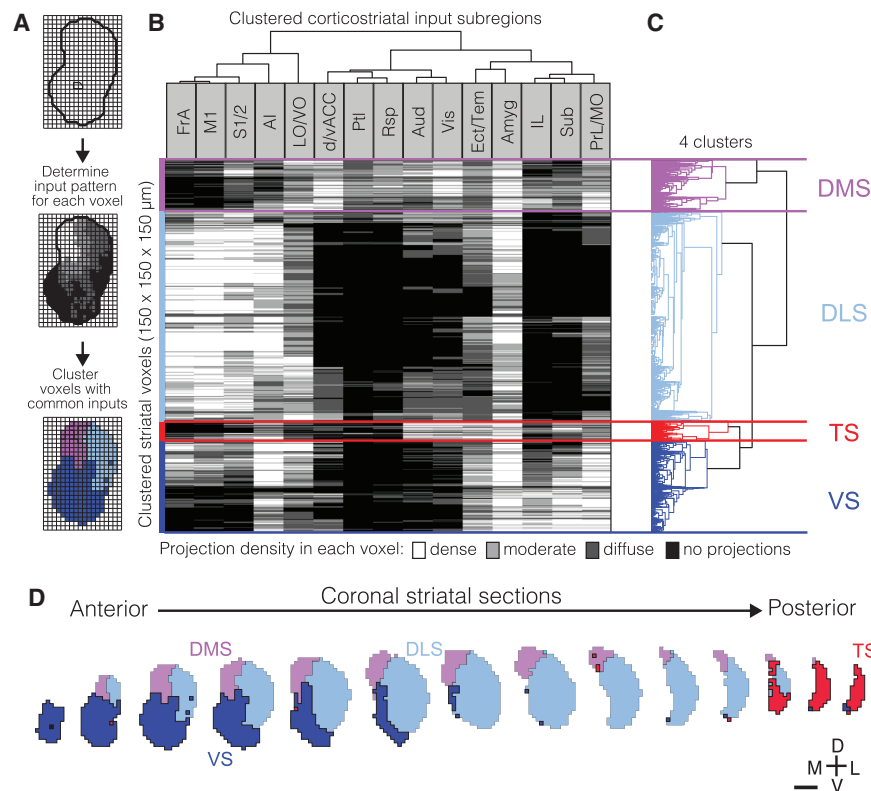
Figure 7. Parcellation of striatum based on cortical inputs. (*A–C*) Unsupervised clustering was used to classify striatal regions (voxels) based on patterns of cortical inputs. (*D*) The four major areas that emerged from the analysis. (VS) Ventral striatum, (DMS) dorsomedial striatum, (DLS) dorsolateral striatum, (TS) tail of striatum, (D) dorsal, (V) ventral, (M) medial, (L) lateral. (Adapted from Hunnicutt et al. 2016.)
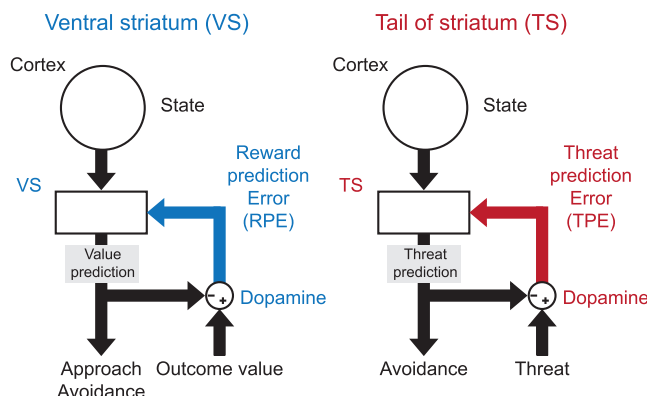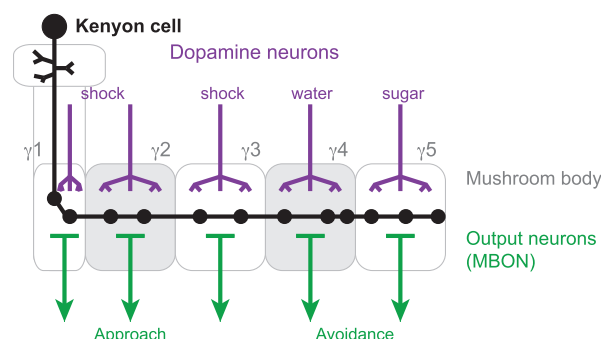
gion. The posterior "tail" of the striatum (TS) lies at the most posterior part of the striatum and receives input from the auditory and visual cortices while it is devoid of input from the motor and somatosensory cortices. It has been regarded as the "sensory" striatum (Xiong et al. 2015; Guo et al. 2018). In recent studies that sought to partition the striatum using unsupervised classification based on their cortical inputs, the TS emerged as the fourth region that receives a unique set of inputs in addition to the three commonly used regions (ventral, dorsomedial, and dorsolateral) (Fig. 7; Hintiryan et al. 2016; Hunnicutt et al. 2016). Furthermore, these four striatal regions receive distinct input from the thalamus (Hunnicutt et al. 2016). Auditory and visual stimuli can signal potential threats from a remote source. The TS is well positioned to detect and modulate these sensory inputs predicting potential threats. Dopamine signals encoding threat or novelty may assist associative learning or boost information flow from the TS to its downstream structures.

COMPARTMENTALIZATION OF STRIATUM BASED ON DOPAMINE SIGNALS

Increasing evidence, including that discussed above, indicates that different regions of the striatum receive distinct dopamine signals. Dopamine signals activated by reward are widespread in the striatum (Menegas et al. 2017) but devoid in some areas such as TS and some

part of DS (Menegas et al. 2017, 2018; Howe and Dombeck 2016). Dopamine signals in the DS tend to be modulated also by an animal's movement or choice (Howe and Dombeck 2016; Parker et al. 2016; Kim and Uchida 2017) (also see relevant data from SNc; Jin and Costa 2010; Barter et al. 2015; Kim et al. 2015). Complementary to these studies are recent findings demonstrating distinct effects of optogenetic stimulation of dopamine subpopulations (Howe and Dombeck 2016; da Silva et al. 2018; Keiflin et al. 2018; Menegas et al. 2018; Saunders et al. 2018).

An emerging view from these recent results is that the striatum is parceled into multiple compartments with each receiving a unique set of dopamine signals. At present, it remains unclear how many compartments exist and whether there are clear boundaries (i.e., whether these compartments are truly discrete or dopamine signals gradually change throughout the striatum). Some previous studies have hinted discrete boundaries. Dopamine axons projecting to the TS are constrained within this region and do not spread into more anterior areas such as the DLS. Furthermore, the unique response properties of certain dopamine axons—strong activation by external threats and novelty but negligible activation by reward—is again restricted to the TS (Menegas et al. 2018). Although other dopamine neuron populations project to more or less restricted areas in the striatum and different responses were observed across areas, whether dopamine axons are truly

A **Algorithmic model: Two axes of reinforcement learning by dopamine**
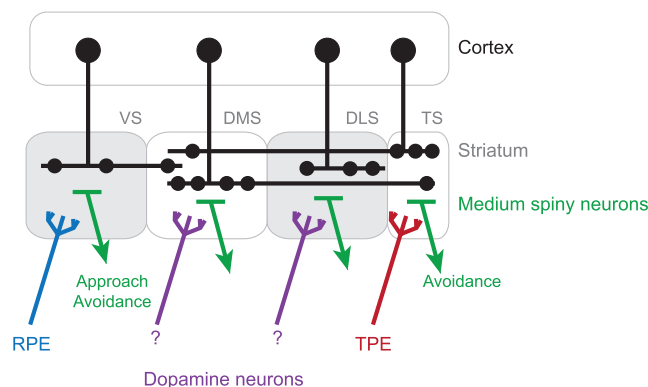
Figure 8. Models of reinforcement learning circuits. (*A*) Algorithmic models. There are at least two types of reinforcement learning using dopamine in mammals. (*B*) Anatomical model of the insect dopamine-mushroom body circuit (after Waddell 2016). Note that in *Drosophila*, it is thought that dopamine suppresses MBON activity and resulting behavior via synaptic depression. (*C*) Anatomical model of the mouse dopamine-striatum circuit. (*B*, Modified from Waddell 2016, with permission from Elsevier.)

segregated in a compartmental fashion remains to be clarified (Beier et al. 2015; Lerner et al. 2015; de Jong et al. 2019). Of note, a recent study found that there is a small patch of the ventral striatum (ventromedial shell) that receives dopamine input which is activated by unexpected aversive outcomes and to cues that predict them (de Jong et al. 2019).

A compartmental organization of a dopamine-recipient area was first noticed in the fruit fly, *Drosophila* (Owald and Waddell 2015; Cohn et al. 2015). In this species, the

area called mushroom body receives dopamine inputs. The mushroom body is parceled into several compartments that are innervated by dopamine input encoding different signals such as punishment and reward (Fig. 8B). Interestingly, some of these compartments correspond to different kinds of reward (water, sweet, or nutrient) or short- versus long-term memory of reward (Burke et al. 2012; Liu et al. 2012; Huetteroth et al. 2015; Yamagata et al. 2015). Axons of Kenyon cells, which convey olfactory information, travel through several compart-

ments—that is, a single axon conveying sensory information can be modulated by different dopamine signals at each compartment in a relatively independent manner (Cohn et al. 2015). Subsequently, neurons in the mushroom body (mushroom body output neurons, MBON) in each compartment control corresponding behaviors such as approach and avoidance (Fig. 8B).

It is noteworthy that although neurons in a given cortical area tend to project to a specific striatal region, some cortical areas arborize across multiple subareas, often longitudinally in the striatum in mammals (Fig. 8C; Selemon and Goldman-Rakic 1985; Mailly et al. 2013; Seger 2013; Hunnicutt et al. 2016). For instance, neurons in the auditory and visual cortices project mainly to the TS but some axons project more anteriorly, with some reaching all the way to the anterior part of the DMS (Hunnicutt et al. 2016). These anatomical observations indicate that axons of single cortical neurons may interact with multiple compartments in the striatum receiving different dopamine signals even in mammals. Thus, across different species, dopamine systems may have remarkable similarities in the basic organization and the modes of functioning.

CONCLUSION

Recent studies have provided firm evidence supporting the diversity of dopamine neurons. Importantly, dopamine neurons projecting to different targets signal distinct information, and regulate different aspects of behavior. An emerging picture from these studies is that different compartments of the dopamine-recipient area—striatum in mammals and mushroom body in fruit flies—control different needs of the animal such as approaching to reward and avoiding dangers supported by distinct associative memories.

Compared to the VS (canonical dopamine system) and TS (threat learning system), functions of dopamine in the DS (dorsomedial and DLS) are less understood theoretically. At a glance, the existing data on dopamine in the DS appear to indicate that it requires different conceptual and theoretical frameworks than reinforcement learning. Given that functions of dopamine in VS and TS can be understood in the framework of prediction error-based reinforcement learning, however, it would be worthwhile to explore the possibility that functions of other parts of the striatum—dorsomedial and DLS—can also be understood similarly in the framework of reinforcement learning (Nakahara 2014; Lau et al. 2017; Gardner et al. 2018). Another important question is how the activities of neurons in different compartments are coordinated or integrated to ultimately trigger a coherent behavior to resolve conflicts between different needs ("weal and woe"). Gathering and integrating new findings on diversity of dopamine neurons will be indispensable to truly understand the function of dopamine systems in adaptive behaviors.

ACKNOWLEDGMENTS

REFERENCES

Babayan BM, Uchida N, Gershman SJ. 2018. Belief state representation in the dopamine system. *Nat Commun* **9:** 1891. doi:10.1038/s41467-018-04397-0

Barter JW, Li S, Lu D, Bartholomew RA, Rossi MA, Shoemaker CT, Salas-Meza D, Gaidis E, Yin HH. 2015. Beyond reward prediction errors: the role of dopamine in movement kinematics. *Front Integr Neurosci* **9:** 39. doi:10.3389/fnint.2015.00039

Bayer HM, Glimcher PW. 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47:** 129–141. doi:10.1016/j.neuron.2005.05.020

Beier KT, Steinberg EE, DeLoach KE, Xie S, Miyamichi K, Schwarz L, Gao XJ, Kremer EJ, Malenka RC, Luo L. 2015. Circuit architecture of VTA dopamine neurons revealed by systematic input-output mapping. *Cell* **162:** 622–634. doi:10.1016/j.cell.2015.07.015

Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K. 2005. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat Neurosci* **8:** 1263–1268. doi:10.1038/nn1525

Bromberg-Martin ES, Matsumoto M, Hong S, Hikosaka O. 2010a. A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J Neurophysiol* **104:** 1068–1076. doi:10.1152/jn.00158.2010

Bromberg-Martin ES, Matsumoto M, Nakahara H, Hikosaka O. 2010b. Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* **67:** 499–510. doi:10.1016/j.neuron.2010.06.031

Brown HD, McCutcheon JE, Cone JJ, Ragozzino ME, Roitman MF. 2011. Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. *Eur J Neurosci* **34:** 1997–2006. doi:10.1111/j.1460-9568.2011.07914.x

Burke CJ, Huetteroth W, Owald D, Perisse E, Krashes MJ, Das G, Gohl D, Silies M, Certel S, Waddell S. 2012. Layered reward signalling through octopamine and dopamine in *Drosophila*. *Nature* **492:** 433–437. doi:10.1038/nature11614

Chung K, Wallace J, Kim S-Y, Kalyanasundaram S, Andalman AS, Davidson TJ, Mirzabekov JJ, Zalocusky KA, Mattis J, Denisin AK, et al. 2013. Structural and molecular interrogation of intact biological systems. *Nature* **497:** 332–337. doi:10.1038/nature12107

Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482:** 85–88. doi:10.1038/nature10754

Cohen JY, Amoroso MW, Uchida N. 2015. Serotonergic neurons signal reward and punishment on multiple timescales. *Elife* **4:** e06346. doi: 10.7554/eLife.06346

Cohn R, Morantte I, Ruta V. 2015. Coordinated and compartmentalized neuromodulation shapes sensory processing in *Drosophila*. *Cell* **163:** 1742–1755. doi:10.1016/j.cell.2015.11.019

da Silva JA, Tecuapetla F, Paixão V, Costa RM. 2018. Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* **554:** 244–248. doi:10.1038/nature25457

de Jong JW, Afjei SA, Pollak Dorocic I, Peck JR, Liu C, Kim CK, Tian L, Deisseroth K, Lammel S. 2019. A neural circuit mech-

anism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron* **101:** 133–151.e7. doi:10.1016/j.neuron.2018.11.005

Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge S, Daw N, Tank D, et al. 2018. Specialized and spatially organized coding of sensory, motor, and cognitive variables in midbrain dopamine neurons. bioRxiv doi:10.1101/456194

Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. 2015. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* **525:** 243–246. doi:10.1038/nature14855

Eshel N, Tian J, Bukwich M, Uchida N. 2016. Dopamine neurons share common response function for reward prediction error. *Nat Neurosci* **19:** 479–486. doi:10.1038/nn.4239

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PEM, Akil H. 2011. A selective role for dopamine in stimulus–reward learning. *Nature* **469:** 53–57. doi:10.1038/nature09588

Gardner MPH, Schoenbaum G, Gershman SJ. 2018. Rethinking dopamine as generalized prediction error. *Proc Biol Sci* **285:** 20181645. doi:10.1098/rspb.2018.1645

Grimm J, Mueller A, Hefti F, Rosenthal A. 2004. Molecular basis for catecholaminergic neuron diversity. *Proc Natl Acad Sci* **101:** 13891–13896. doi:10.1073/pnas.0405340101

Gunaydin LA, Grosenick L, Finkelstein JC, Kauvar IV, Fenno LE, Adhikari A, Lammel S, Mirzabekov JJ, Airan RD, Zalocusky KA, et al. 2014. Natural neural projection dynamics underlying social behavior. *Cell* **157:** 1535–1551. doi:10.1016/j.cell.2014.05.017

Guo L, Walker WI, Ponvert ND, Penix PL, Jaramillo S. 2018. Stable representation of sounds in the posterior striatum during flexible auditory decisions. *Nat Commun* **9:** 1534. doi:10.1038/s41467-018-03994-3

Hart AS, Rutledge RB, Glimcher PW, Phillips PEM. 2014. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *J Neurosci* **34:** 698–704. doi:10.1523/JNEUROSCI.2489-13.2014

Hintiryan H, Foster NN, Bowman I, Bay M, Song MY, Gou L, Yamashita S, Bienkowski MS, Zingg B, Zhu M, et al. 2016. The mouse cortico-striatal projectome. *Nat Neurosci* **19:** 1100–1114. doi:10.1038/nn.4332

Horvitz JC, Stewart T, Jacobs BL. 1997. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res* **759:** 251–258. doi:10.1016/S0006-8993(97)00265-5

Howe MW, Dombeck DA. 2016. Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature* **535:** 505–510. doi:10.1038/nature18942

Huetteroth W, Perisse E, Lin S, Klappenbach M, Burke C, Waddell S. 2015. Sweet taste and nutrient value subdivide rewarding dopaminergic neurons in *Drosophila*. *Curr Biol* **25:** 751–758. doi:10.1016/j.cub.2015.01.036

Hunnicutt BJ, Jongbloets BC, Birdsong WT, Gertz KJ, Zhong H, Mao T. 2016. A comprehensive excitatory input map of the striatum reveals novel functional organization. *Elife* **5:** e19103. doi:10.7554/eLife.19103

James W. 1890. *The principles of psychology.* Henry Holt and Company, New York.

Jin X, Costa RM. 2010. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466:** 457–462. doi:10.1038/nature09263

Kakade S, Dayan P. 2002. Dopamine: generalization and bonuses. *Neural Netw* **15:** 549–559. doi:10.1016/S0893-6080(02)00048-5

Kamin L. 1969. Selective association and conditioning. In *Fundamental issues in associative learning: proceedings of a symposium held at Dalhousie University, Halifax, June 1968* (ed. Mackintosh NJ, Honig WK), pp. 42–64. Dalhousie University Press, Halifax.

Keiflin R, Pribut HJ, Shah NB, Janak PH. 2018. Ventral tegmental dopamine neurons participate in reward identity predictions. *Curr Biol* **29:** 93–103. doi:10.1016/j.cub.2018.11.050

Keller PJ, Schmidt AD, Santella A, Khairy K, Bao Z, Wittbrodt J, Stelzer EHK. 2010. Fast, high-contrast imaging of animal development with scanned light sheet–based structured-illumination microscopy. *Nat Methods* **7:** 637–642. doi:10.1038/nmeth.1476

Kim H, Uchida N. 2017. Phasic dopamine response in the dorsal striatum correlates with specific movements. http://www.abstractsonline.com/pp8/index.html#!/4376/presentation/3977.

Kim HF, Ghazizadeh A, Hikosaka O. 2015. Dopamine neurons encoding long-term memory of object value for habitual behavior. *Cell* **163:** 1165–1175. doi:10.1016/j.cell.2015.10.063

Kim CK, Yang SJ, Pichamoorthy N, Young NP, Kauvar I, Jennings JH, Lerner TN, Berndt A, Lee SY, Ramakrishnan C, et al. 2016. Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nat Methods* **13:** 325–328. doi:10.1038/nmeth.3770

Kudo Y, Akita K, Nakamura T, Ogura A, Makino T, Tamagawa A, Ozaki K, Miyakawa A. 1992. A single optical fiber fluorometric device for measurement of intracellular Ca^{2+} concentration: its application to hippocampal neurons in vitro and in vivo. *Neuroscience* **50:** 619–625. doi:10.1016/0306-4522(92)90451-7

Lacey MG, Mercuri NB, North RA. 1989. Two cell types in rat substantia nigra zona compacta distinguished by membrane properties and the actions of dopamine and opioids. *J Neurosci* **9:** 1233–1241. doi:10.1523/JNEUROSCI.09-04-01233.1989

Lak A, Stauffer WR, Schultz W. 2016. Dopamine neurons learn relative chosen value from probabilistic rewards. *Elife* **5:** e18044. doi:10.7554/eLife.18044

Lammel S, Hetzel A, Häckel O, Jones I, Liss B, Roeper J. 2008. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57:** 760–773. doi:10.1016/j.neuron.2008.01.022

Lau B, Monteiro T, Paton JJ. 2017. The many worlds hypothesis of dopamine prediction error: implications of a parallel circuit architecture in the basal ganglia. *Curr Opin Neurobiol* **46:** 241–247. doi:10.1016/j.conb.2017.08.015

Lerner TN, Shilyansky C, Davidson TJ, Evans KE, Beier KT, Zalocusky KA, Crow AK, Malenka RC, Luo L, Tomer R, et al. 2015. Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* **162:** 635–647. doi:10.1016/j.cell.2015.07.014

Lima SQ, Hromádka T, Znamenskiy P, Zador AM. 2009. PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS One* **4:** e6099. doi:10.1371/journal.pone.0006099

Liu C, Plaçais P-Y, Yamagata N, Pfeiffer BD, Aso Y, Friedrich AB, Siwanowicz I, Rubin GM, Preat T, Tanimoto H. 2012. A subset of dopamine neurons signals reward for odour memory in *Drosophila*. *Nature* **488:** 512–516. doi:10.1038/nature11304

Ljungberg T, Apicella P, Schultz W. 1992. Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* **67:** 145–163. doi:10.1152/jn.1992.67.1.145

Mailly P, Aliane V, Groenewegen HJ, Haber SN, Deniau J-M. 2013. The rat prefrontostriatal system analyzed in 3D: evidence for multiple interacting functional units. *J Neurosci* **33:** 5718–5727. doi:10.1523/JNEUROSCI.5248-12.2013

Matias S, Lottem E, Dugué GP, Mainen ZF. 2017. Activity patterns of serotonin neurons underlying cognitive flexibility. *Elife* **6:** e20552. doi:10.7554/eLife.20552

Matsumoto M, Hikosaka O. 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459:** 837–841. doi:10.1038/nature08028

Matsumoto H, Tian J, Uchida N, Watabe-Uchida M. 2016. Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *Elife* **5:** e17328. doi:10.7554/eLife.17328

Menegas W, Bergan JF, Ogawa SK, Isogai Y, Umadevi Venkataraju K, Osten P, Uchida N, Watabe-Uchida M. 2015. Dopamine neurons projecting to the posterior striatum form an anatomi-

cally distinct subclass. *Elife* **4:** e10032. doi:10.7554/eLife.10032

Menegas W, Babayan BM, Uchida N, Watabe-Uchida M. 2017. Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *Elife* **6:** e21886. doi:10.7554/eLife.21886

Menegas W, Akiti K, Amo R, Uchida N, Watabe-Uchida M. 2018. Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nat Neurosci* **21:** 1421–1430. doi:10.1038/s41593-018-0222-1

Mirenowicz J, Schultz W. 1994. Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* **72:** 1024–1027. doi:10.1152/jn.1994.72.2.1024

Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, et al. 2015. Human-level control through deep reinforcement learning. *Nature* **518:** 529–533. doi:10.1038/nature14236

Nakahara H. 2014. Multiplexing signals in reinforcement learning with internal models and dopamine. *Curr Opin Neurobiol* **25:** 123–129. doi:10.1016/j.conb.2014.01.001

Olds J, Milner P. 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J Comp Physiol Psychol* **47:** 419–427. doi:10.1037/h0058775

Oleson EB, Gentry RN, Chioma VC, Cheer JF. 2012. Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J Neurosci* **32:** 14804–14808. doi:10.1523/JNEUROSCI.3087-12.2012

Owald D, Waddell S. 2015. Olfactory learning skews mushroom body output pathways to steer behavioral choice in *Drosophila*. *Curr Opin Neurobiol* **35:** 178–184. doi:10.1016/j.conb.2015.10.002

Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND, Witten IB. 2016. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nat Neurosci* **19:** 845–854. doi:10.1038/nn.4287

Poulin J-F, Zou J, Drouin-Ouellet J, Kim K-YA, Cicchetti F, Awatramani RB. 2014. Defining midbrain dopaminergic neuron diversity by single-cell gene expression profiling. *Cell Rep* **9:** 930–943. doi:10.1016/j.celrep.2014.10.008

Rebec GV, Christensen JR, Guerra C, Bardo MT. 1997. Regional and temporal differences in real-time dopamine efflux in the nucleus accumbens during free-choice novelty. *Brain Res* **776:** 61–67. doi:10.1016/S0006-8993(97)01004-4

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning II: current research and theory* (ed. Black AH, Prokasy WF), pp. 64–99. Appleton-Century-Crofts, New York.

Roeper J. 2013. Dissecting the diversity of midbrain dopamine neurons. *Trends Neurosci* **36:** 336–342. doi:10.1016/j.tins.2013.03.003

Roitman MF, Wheeler RA, Wightman RM, Carelli RM. 2008. Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nat Neurosci* **11:** 1376–1377. doi:10.1038/nn.2219

Saunders BT, Richard JM, Margolis EB, Janak PH. 2018. Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nat Neurosci* **21:** 1072–1083. doi:10.1038/s41593-018-0191-4

Schultz W. 2013. Updating dopamine reward signals. *Curr Opin Neurobiol* **23:** 229–238. doi:10.1016/j.conb.2012.11.012

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science* **275:** 1593–1599. doi:10.1126/science.275.5306.1593

Seger CA. 2013. The visual corticostriatal loop through the tail of the caudate: circuitry and function. *Front Syst Neurosci* **7:** 104. doi:10.3389/fnsys.2013.00104

Selemon LD, Goldman-Rakic PS. 1985. Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *J Neurosci* **5:** 776–794. doi:10.1523/JNEUROSCI.05-03-00776.1985

Starkweather CK, Babayan BM, Uchida N, Gershman SJ. 2017. Dopamine reward prediction errors reflect hidden-state inference across time. *Nat Neurosci* **20:** 581–589. doi:10.1038/nn.4520

Starkweather CK, Gershman SJ, Uchida N. 2018. The medial prefrontal cortex shapes dopamine reward prediction errors under state uncertainty. *Neuron* **98:** 616–629.e6. doi:10.1016/j.neuron.2018.03.036

Steinberg EE, Janak PH. 2013. Establishing causality for dopamine in neural function and behavior with optogenetics. *Brain Res* **1511:** 46–64. doi:10.1016/j.brainres.2012.09.036

Steinfels GF, Heym J, Strecker RE, Jacobs BL. 1983. Behavioral correlates of dopaminergic unit activity in freely moving cats. *Brain Res* **258:** 217–228. doi:10.1016/0006-8993(83)91145-9

Stuber GD, Klanker M, de Ridder B, Bowers MS, Joosten RN, Feenstra MG, Bonci A. 2008. Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons. *Science* **321:** 1690–1692. doi:10.1126/science.1160873

Sutton RS, Barto AG. 1998. *Reinforcement learning: an introduction*. MIT Press, Cambridge, MA.

Tesauro G. 1995. Temporal difference learning and TD-Gammon. *Commun ACM* **38:** 58–68. doi:10.1145/203330.203343

Tian J, Uchida N. 2015. Habenula lesions reveal that multiple mechanisms underlie dopamine prediction errors. *Neuron* **87:** 1304–1316. doi:10.1016/j.neuron.2015.08.028

Tsai H-C, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K. 2009. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324:** 1080–1084. doi:10.1126/science.1168878

Ungless MA, Grace AA. 2012. Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. *Trends Neurosci* **35:** 422–430. doi:10.1016/j.tins.2012.02.003

Waddell S. 2016. Neural plasticity: dopamine tunes the mushroom body output network. *Curr Biol* **26:** R109–R112. doi:10.1016/j.cub.2015.12.023

Watabe-Uchida M, Zhu L, Ogawa SK, Vamanrao A, Uchida N. 2012. Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron* **74:** 858–873. doi:10.1016/j.neuron.2012.03.017

Watabe-Uchida M, Eshel N, Uchida N. 2017. Neural circuitry of reward prediction error. *Annu Rev Neurosci* **40:** 373–394. doi:10.1146/annurev-neuro-072116-031109

Watkins CJCH, Dayan P. 1992. *Q*-learning. *Mach Learn* **8:** 279–292.

Wickersham IR, Lyon DC, Barnard RJO, Mori T, Finke S, Conzelmann K-K, Young JAT, Callaway EM. 2007. Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron* **53:** 639–647. doi:10.1016/j.neuron.2007.01.033

Wise RA. 2004. Dopamine, learning and motivation. *Nat Rev Neurosci* **5:** 483–494. doi:10.1038/nrn1406

Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C, et al. 2011. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72:** 721–733. doi:10.1016/j.neuron.2011.10.028

Xiong Q, Znamenskiy P, Zador AM. 2015. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* **521:** 348–351. doi:10.1038/nature14225

Yamagata N, Ichinose T, Aso Y, Plaçais P-Y, Friedrich AB, Sima RJ, Preat T, Rubin GM, Tanimoto H. 2015. Distinct dopamine neurons mediate reward signals for short- and long-term memories. *Proc Natl Acad Sci* **112:** 578–583. doi:10.1073/pnas.1421930112

Cold Spring Harbor Symposia on Quantitative Biology

Multiple Dopamine Systems: Weal and Woe of Dopamine

Mitsuko Watabe-Uchida and Naoshige Uchida

References	This article cites 85 articles, 11 of which can be accessed free at: **http://symposium.cshlp.org/content/83/83.full.html#ref-list-1**
Creative Commons License	This article is distributed under the terms of the http://creativecommons.org/licenses/by-nc/4.0/, which permits reuse and redistribution, except for commercial purposes, provided that the original author and source are credited.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**