Cell PRESS

Special Issue: Hippocampus and Memory

The hippocampal–striatal axis in learning, prediction and goal-directed behavior

C.M.A. Pennartz[1], R. Ito[2,3], P.F.M.J. Verschure[4,5], F.P. Battaglia[1] and T.W. Robbins[6]

[1] Cognitive and Systems Neuroscience Group, Swammerdam Institute for Life Sciences, Center for Neuroscience, Sciencepark 904, 1098 XH, Amsterdam, The Netherlands
[2] Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, UK
[3] Department of Psychology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON, M1C 1A4, Canada
[4] Laboratory of Synthetic Perceptive Emotive and Cognitive Systems (SPECS), Department of Technology, Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Spain
[5] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[6] Behavioral and Clinical Neuroscience Institute and Department of Experimental Psychology, University of Cambridge, Cambridge, Downing Street CB2 3EB, UK

The hippocampal formation and striatum subserve declarative and procedural memory, respectively. However, experimental evidence suggests that the ventral striatum, as opposed to the dorsal striatum, does not lend itself to being part of either system. Instead, it may constitute a system integrating inputs from the amygdala, prefrontal cortex and hippocampus to generate motivational, outcome-predicting signals that invigorate goal-directed behaviors. Inspired by reinforcement learning models, we suggest an alternative scheme for computational functions of the striatum. Dorsal and ventral striatum are proposed to compute outcome predictions largely in parallel, using different types of information as input. The nature of the inputs to striatum is furthermore combinatorial, and the specificity of predictions transcends the level of scalar value signals, incorporating episodic information.

Introduction

Distinct forms of memory are considered to be mediated by different brain systems. Traditionally, a dichotomy is applied between declarative (explicit) memory versus non-declarative (procedural, implicit) memory [1]. Declarative memory refers to our ability to recall events from the past deliberately and consciously; procedural memory refers to motor or cognitive skills that come to be executed automatically and are recalled unconsciously. Strong evidence implicates the hippocampal–temporal lobe system in declarative memory and the striatum and connected basal ganglia structures in procedural learning and habit formation [2–4]. The exact functions of the hippocampus (HPC) are far from clear, but the weight of evidence favors a role in episodic memory, which stores information about individually experienced events, set in a specific spatio-temporal context [5–8]. By contrast, procedural memories are thought to be stored by synaptic modifications in

Glossary

Conditioned place preference (CPP) test: behavioral paradigm assessing reinforcing effects of drugs or rewards. Animals undergo conditioning sessions in different environments (or spatial compartments), only one of which is associated with the drug or reward. The acquisition of a spatial–reward association is indicated by the animal's preference for the environment previously paired with reward (but currently in its absence).

Conditioned reinforcer: a previously neutral stimulus (e.g. stimulus light) that acquires the ability to reinforce behavior upon which it is contingent, by virtue of having been paired predictively with a primary reinforcer (e.g. sucrose).

Matrix and striosomes: in the dorsal striatum, small regions are discerned, called striosomes or patches, which are surrounded by a matrix region. These compartments differ in neurochemical makeup and input/output connectivity. Medium-sized spiny neurons in striosomes have been reported to project to dopamine neurons in SNC and VTA, whereas in the matrix this type of neuron projects to output regions of the basal ganglia, viz. pallidal structures and SNR [13,16]. Neurochemical compartmentalization in VS is more complex [13].

Model-free reinforcement learning: class of RL algorithms in which the association between an organism's state or action and its outcome is cached (i.e. stored, captured) in a scalar signal summarizing its long-term value, without specifying the nature or features of the outcome. This class contrasts with model-based approaches, in which state or action associations with outcome are learned indirectly, by constructing a model of the organism's environment. This model can be high-dimensional, i.e. incorporating many features of states or actions as well as outcomes [75].

Motivation: a state of desire or energy to carry out a certain action, triggered by intrinsic and extrinsic factors, which can be aversive or appetitive.

Outcome: payoff or consequence of a given stimulus or action, which can be positive (rewarding) or negative (aversive). An outcome is not necessarily identical to a reinforcer, which by definition should alter future behavior related to its presentation.

Pavlovian to instrumental transfer: phenomenon in which a pavlovian conditioned stimulus invigorates (if appetitive) or reduces (if aversive) the rate of an appetitively motivated instrumental behavior (e.g. lever pressing) when it is presented non-contingently during instrumental performance.

Reinforcement learning (RL): type of learning in which an agent initially responds to stimuli (or input states) by trial and error, and learns to improve its responses based on reinforcing feedback from the environment. This reinforcing feedback specifies only how good or bad the agent's action was, not how the agent should have been responding given a certain situation.

Sharp wave-ripple: electrophysiological pattern of activity in the hippocampal electroencephalogram (EEG) characterized by high-frequency (150–250 Hz) waxing-and-waning oscillations (ripples) and steep negative potentials (sharp waves) coupled to strong dendritic depolarization of pyramidal cells.

Striatum: the striatum has been subdivided into two main regions, DS and VS. The boundary between these regions is not well defined, and neuroanatomical studies indicate it is more appropriate to speak of a ventromedial to

Corresponding author: Pennartz, C.M.A. (c.m.a.pennartz@uva.nl).

dorsolateral gradient [13]. DS is also referred to as caudate-putamen and is further subdivided in dorsolateral striatum (DLS, putamen) and dorsomedial striatum (DMS, caudate). The VS is subdivided in a (ventromedial) shell and a (dorsolateral) core, each having distinct anatomical and physiological characteristics [13,19] (Figure 1). The term VS is used here when statements apply both to core and shell, or when previous studies cited did not distinguish between them, and the same applies to DS as comprising DMS and DLS.

neocortical–basal ganglia loops. These loops connect specific neocortical areas unidirectionally to striatal subregions, which project to downstream structures such as the pallidum, ventral tegmental area (VTA), substantia nigra pars reticulata (SNR) and pars compacta (SNC). These areas connect to thalamic nuclei that project back to neocortical areas identical to, or close to, the site of origin.

Various parallel loops have been associated with different types of motor and cognitive function. Oculomotor and somatic motor loops originate in the frontal eye fields and (pre)motor cortices, but cognitive and motivational-affective loops associated with prefrontal cortex (PFC), amygdala and HPC have also been identified [9–11]. The striatum has been subdivided into corresponding regions: whereas the dorsolateral striatum (DLS) mediates stimulus–response learning and habit formation, the dorsomedial striatum (DMS) is associated with cognitive functions and action–outcome learning, and ventral striatum (VS) with motivational and affective processing [2,4,12–17]. The VS occupies a peculiar position in this system, challenging the episodic–procedural dichotomy because it possesses key features of striatum [18] but also receives a strong projection from HPC [19] (Figure 1). Thus, is the VS part of the declarative or procedural memory system? The main goal of this review will be to address this question by conceptualizing how hippocampal input to the VS is integrated with other inputs to govern motivational processes, aided by models of reinforcement learning (RL) (see Glossary). Recent experimental findings in the field, considered together with insights from available computational models, will lead us to propose a revised model of limbic corticostriatal circuitry that goes beyond a classical RL architecture.

Causal roles of HPC and VS in different types of learning and memory

Collectively, the HPC and VS have been implicated in a wealth of behavioral processes (e.g. latent inhibition, attention and anxiety), but this review will focus only on a subset of these, viz. behavioral responses to spatial or contextual and discrete cues. The HPC has been divided along the dorsal–ventral axis, with dorsal HPC preferentially involved in spatial learning and ventral HPC in anxiety-related behavior [20,21]. However, other evidence suggests the ventral and dorsal HPC serve a common role in some forms of learning. Lesions and pharmacological inactivation of both dorsal and ventral HPC impair contextual aversive and appetitive conditioning and context-dependent memory retrieval [22–25] (Figure 2a), whereas ventral HPC lesions also impair fear conditioning to discrete auditory cues, reminiscent of basolateral amygdala (BLA) lesion effects ([21], but see [26] for dorsal HPC involvement in delayed-fear conditioning to auditory cues).
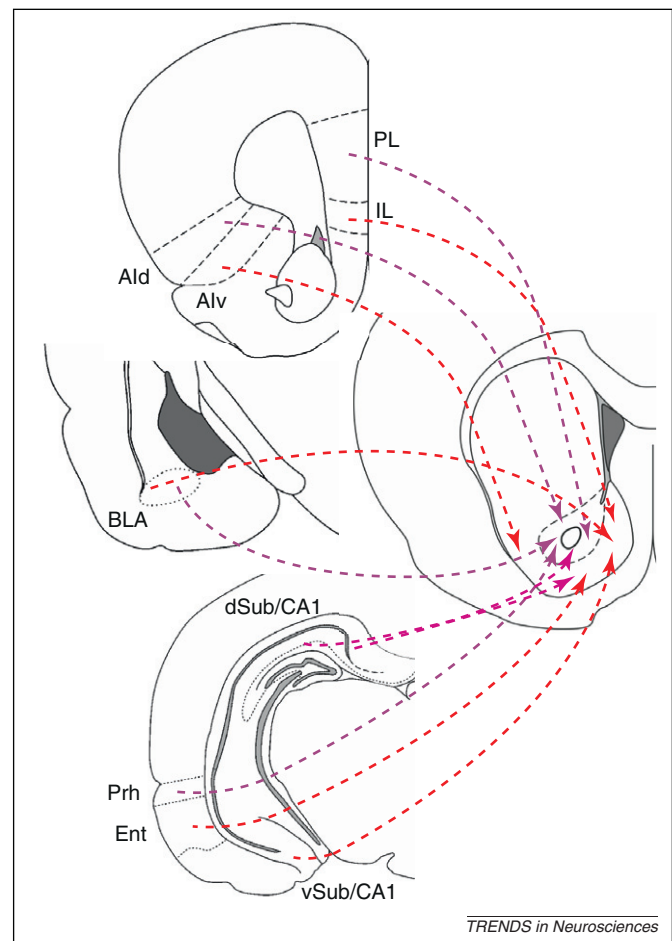

Figure 1. Main cortical and amygdaloid inputs to the rat ventral striatum (VS). Afferent pathways from frontal cortex, basolateral amygdala (BLA), hippocampus (HPC) and adjoining areas are illustrated. Inputs from the midline and intralaminar thalamic nuclei have been left out for simplicity, as well as inputs to the dorsal striatum or striatal elements of the olfactory tubercle. Purple and red arrows indicate projections predominantly reaching the core and shell region of the VS, respectively. Rostrocaudal gradients of innervation are not represented here. Fibers from the ventral subiculum (vSub) and area CA1 reach the medial, ventral and rostral shell, whereas the dorsal subiculum (dSub) and CA1 project primarily to rostral parts of the VS (mixed purple-red). Both in shell and core, these hippocampal inputs converge with inputs from the perirhinal (Prh) and entorhinal (Ent) cortices, BLA and frontal cortex. Abbreviations: AId, dorsal agranular insular cortex; AIv, ventral agranular insular cortex; IL, infralimbic cortex; PL, prelimbic cortex. Sections based on [118].

Thus, dorsal and ventral HPC may subserve qualitatively similar roles in context conditioning, but their contributions may differ according to what constitutes the context representation. A context defined by spatiotemporal cues (a configuration of multiple environmental or idiothetic cues) may predominantly engage dorsal HPC [21], whereas a context defined by non-spatial (e.g. odor, interoceptive and emotional) cues may rely more strongly on ventral HPC [27,28]. This dorsal–ventral distinction is supported, for example, by a decrease in spatial representation and theta rhythm from dorsal to ventral hippocampal area CA3 [29]. However, there is probably considerable overlap in the types of information the two regions process, and the dorsal–ventral divide may be better understood as a functional continuum rather than an absolute division [30–32].

Popularly known as a limbic–motor interface, the VS has been proposed to translate information from HPC into
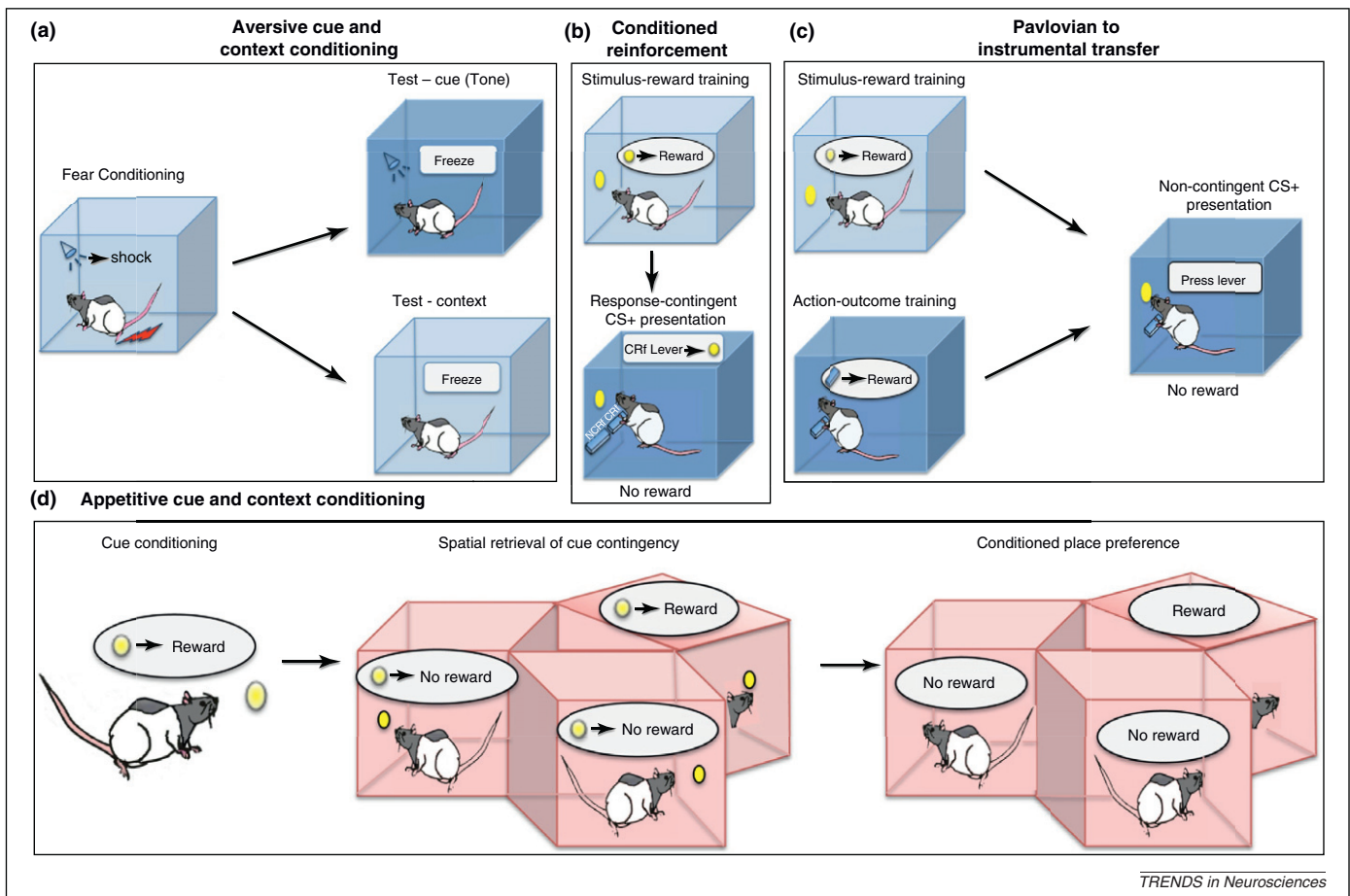
Figure 2. Behavioral tasks that depend on the hippocampus (HPC), amygdala and ventral striatum (VS). **(a)** Aversive cue and context conditioning. In this task, the rat learns that a discrete cue [conditioned stimulus (CS), e.g. tone] and a context in which the training takes place, predict the occurrence of an unconditioned stimulus (US; e.g. an electric shock). Subsequent exposure to the cue or context in the absence of the shock induces freezing behavior (i.e. a conditioned response). **(b)** Conditioned Reinforcement (CRf). In the first phase of training (Stimulus–Reward training), the rat learns that a light cue (CS) predicts reward (e.g. sucrose pellet). In the second phase, the rat learns a new instrumental response (e.g. lever press) for the presentation of a CS on one lever (CRf lever) over another (i.e. non-conditioned reinforcement, NCRf). Ellipse in upper panel symbolizes the acquired association; rectangular box in lower panel denotes behavioral sequence. **(c)** Pavlovian to Instrumental Transfer (PIT). In the first phase, the rat undergoes stimulus–reward (CS–US) training in one environment, and instrumental learning (lever pressing for reward; action–outcome training) in another environment. In the transfer test, the rat receives passive CS presentations during lever pressing in the absence of reward. The VS core and central nucleus of the amygdala are involved in mediating general motivational effects of pavlovian cues on instrumental behavior. However, in a different form of PIT in which different outcomes are associated with two pavlovian stimuli in the stimulus–reward training phase, and two levers in the action–outcome learning stage, the BLA and VS shell support outcome-specific effects of pavlovian cues upon instrumental responses. **(d)** Appetitive cue and context conditioning in the Y maze. The rat initially learns to associate a flashing light cue with sucrose solution. Following acquisition of cue conditioning, the rat needs to learn that the same cue is rewarded only when presented in one chamber of the Y maze in a fixed spatial location (as defined by path integration). Thus, the procedure tests the use of spatial information to retrieve cue contingencies. At the end of the retrieval acquisition, the rat undergoes a conditioned place preference (CPP) test in the absence of reward to assess whether it has developed a preference for the rewarded chamber.

action [33]. Lesion studies showed that the VS is not only important for processing spatial and contextual cues [34–37], but also for BLA-dependent appetitive and aversive cue conditioning [38] and the ability of pavlovian cues to support instrumental responding [39–41]. Apart from the striatal elements of the olfactory tubercle, the VS is commonly differentiated into a core and shell region [13] (Figure 1). Both regions receive input from the BLA, but the shell receives hippocampal input predominantly from ventral CA1 and subiculum, whereas the core receives it from dorsal CA1 and subiculum and from parahippocampal regions [13,19] (Figure 1).

Studies using disconnection lesions have provided evidence for distinct limbic–striatal circuits subserving different forms of conditioning. Evidence indicates a critical role of the HPC–shell pathway in the acquisition of appetitive context conditioning and for retrieval of cue contingencies

based on spatial locations [42] (Figure 2). The function of the sparser HPC–core pathway (Figure 1) is largely unknown, although some evidence supports a role of the core in contextual conditioning and control of spatial behavior [43,44]. By contrast, information transfer between the BLA and core is important for mediating the excitatory effects of pavlovian cues or conditioned reinforcers on behavior [40,45].

Dopamine release in VS plays a key role in mediating affective control over motivated behavior, and its dysregulation may contribute to disorders such as schizophrenia and drug addiction (Box 1). Acute elevation of dopamine concentration in VS, but not dorsal striatum (DS), potentiates effects of conditioned reinforcers on lever pressing [39,46]. This effect is attenuated by BLA lesions, indicating the importance of BLA–dopamine interactions in VS for mediating the effect of reward-predicting stimuli on action

Box 1. Psychopathology of the VS and implications for neuropsychiatric disorders

The involvement of the VS in the reinforcing effects of drugs such as cocaine, nicotine and heroin has led to this structure being linked to drug addiction [120]. There is considerable preclinical evidence to support a role for the VS in drug-seeking behavior in experimental animals [14], and during acute cocaine infusions in humans [121], presumably because of its mediation of motivational effects of conditioned stimuli associated with the drug leading to its anticipation, as well as the unconditioned effects of the drug itself. However, it is less clear that the VS has a major role to play in drug addiction *per se* beyond the initiation of drug abuse [14,122]. The concept of a transition of neural control over drug-seeking behavior from the VS to DS [14] is in fact consistent with the present hypothesis that the VS provides an interface between declarative and procedural or habit-based learning.

Functional neuroimaging studies in alcoholics [123], compulsive gamblers [124] and patients with attention deficit hyperactivity disorder [125] (although not depression [126]), suggest that under-activation of the VS may be associated with impulsive behavioral tendencies, which may arise from a dysregulation of anticipatory tendencies to conditioned stimuli. Moreover, apparent overdosing of the mesolimbic dopaminergic pathway in Parkinson's disease (PD) by dopaminergic medication can lead to impaired inhibitory control associated with compulsive gambling in PD patients (reviewed in [127]).

Psychotic symptoms in schizophrenia have also been associated with VS dysfunction. It was assumed for many years that antipsychotic effects of drugs such as haloperidol were exerted via effects on the mesolimbic dopamine system. This system was assumed to be overactive, producing aberrant 'incentive salience' in response to environmental stimuli (presumably both cues and contexts) and leading to delusional phenomena [128]. This is consistent with the anatomical connectivity of the HPC, which is known to be affected early in the course of schizophrenia [129]. However, recent evidence [130] suggests that the main striatal region exhibiting dopamine overactivity in schizophrenia is the caudate nucleus (the DMS in rats) rather than the VS, and this apparent mismatch has yet to be resolved.

[39]. Recently, selective dopamine elevation by direct d-amphetamine infusions in the shell of the VS was shown to enhance HPC-dependent control over conditioned place preference (CPP), whereas in the core, this treatment attenuated HPC control over this form of learning [47]. Such findings demonstrate regional differences within the VS in the way dopamine regulates limbic information processing [47].

Taken together, lesion and pharmacological evidence support the existence of distinct limbic corticostriatal loops involved in processing different types of associative information. The hippocampal–VS (shell) pathway is critical for associating contextual–positional information with outcomes and the BLA–VS (core) pathway for discrete cue–outcome associations. Moreover, dopamine selectively modulates the strength or gain of associative control over motivated behavior in a regionally specific manner.

Neural coding of different types of information in the HPC and VS

In vivo recordings in freely behaving animals have provided insights into how the HPC and striatum encode information on context and motivation at high temporal resolution. Following the discovery of hippocampal place cells in area CA1, which fire specifically when an animal occupies a particular location in an open environment [48], further studies in tasks that differed from open-space exploration indicated that behavioral variables other than place can also be coded by hippocampal neurons, including sensory cues [49,50] and sequential–temporal aspects of behavioral episodes [51]. In agreement with its role in episodic memory [5–7], we will therefore refer to the nature of hippocampal CA1 output as spatial–episodic.

Aside from the subiculum and area CA1 [19], perirhinal and entorhinal cortex provide significant inputs to VS [52,53]. Subicular neurons are sensitive to an animal's location, albeit less specifically than found in area CA1 [54,55]. The medial entorhinal cortex is thought to encode positional information by way of grid cells, as well as head direction information [56]. By contrast, perirhinal and lateral entorhinal projections probably convey object-related information with low spatial specificity [57]. Altogether, the

HPC and its adjoining areas supply the VS with a rich stream of information regarding the animal's position and orientation in geometric space as well as relevant objects and temporal context (Box 2).

Neurons in rodent and primate VS respond to all behavioral elements of goal-directed sequences relevant to an ongoing task, even when associated with aversive outcome [58–64]. Typically, the entire task sequence, including outcome, is tessellated by VS firing patterns (Figure 3). For example, it has been observed that VS neurons generate diverse firing responses to both aversive and appetitive outcomes, in combination with oromotor responses, in a classical conditioning task [65]. Moreover, a learning-dependent development of responses to auditory stimuli predicting the outcome was observed [65]. In a reward-seeking task on a triangular track marked by three sites distinguished by qualitatively different rewards, VS neurons fired selectively in anticipation or during delivery of reward [62] (Figure 3), in agreement with outcome-specific coding observed in other studies [59,66,67].

Firing patterns in the VS are generally marked by a strong motivational component. In a cued arm-reaching movement task in monkeys, VS neurons expressed sustained increments in firing rate, which occurred regardless of whether an arm movement was made, and were thus inferred to reflect reward expectancy [68]. Sustained expectancy-related firing patterns, such as firing-rate 'ramps', have also been observed in other studies [60,62,69]. These ramps can take into account the place or identity of expected rewards [62] (Figure 3b,c). Expectancy signals, generated in association with cues and movements, are sculpted during learning and are sensitive to changes in outcome contingencies, confirming their motivational nature [59,62,63,65,68]. VS neurons not only fire in anticipation of outcomes, but subsets may also respond during reward consumption. Thus, whereas the HPC codes the context framing episodic experience, VS neural coding, albeit highly versatile and complex (Box 3), is centered on the relevant elements of goal-directed tasks in conjunction with their motivational component (i.e. the extent to which stimuli, context and actions predict outcome). Whereas it is not always clear whether the influence of motivational factors involving the VS is due to

Box 2. Dynamics of hippocampal–striatal communication during behavior and sleep

Behavioral results using CPP paradigms (see Figure 2 and main text) have clarified how the HPC–VS axis may serve as a model system for studying how brain structures communicate during behavior in general. How is this communication instantiated, and which mechanisms mediate synaptic plasticity in this system? During active behavior, mass activity in the HPC is characterized by a theta rhythm (i.e. 6–12 Hz) in local field potentials. VS neurons often fire preferentially within a narrow phase range defined by the hippocampal theta cycle, suggesting a coupling between these systems during goal-directed behavior [112,131]. The firing of a HPC place cell in a rat traversing the corresponding place field shows a progressively earlier and earlier phase in the theta cycle, a phenomenon called phase precession [132]. At an ensemble level, phase precession enables a time-compressed representation of a sequence of subsequently visited places [133]. VS cells exhibiting an anticipatory ramp in firing rate showed precession to hippocampal theta oscillations as well [69], suggesting that reward-related signals are temporally aligned with spatial–episodic information during anticipation and possibly decision-making. A special mode of hippocampal processing, found in area CA3, may be the forward sweep [134], which is expressed as a rapid sequence of ensemble activity coding for places ahead of an animal's actual position. During forward sweeps, VS cells also fire [135], which may indicate a mechanism for deliberating about goal prospects before committing to a behavioral choice.

Another mode of communication appears during sleep, in particular when neocortical slow waves appear. In the HPC, SWS is marked by irregular EEG activity and intermittent sharp wave-ripple activity (150–250 Hz) [48,136]. During ripples occurring in post-experiential sleep, the HPC replays firing patterns characteristic of the preceding behavioral experience, and significantly more so than during sleep prior to this experience [137]. Because disruption of ripple activity impairs spatial learning [138], memory consolidation is probably benefiting from this process. Also during SWS the hippocampal formation communicates intensively with the VS, as suggested by the modulation of VS firing rates by ripples [139]. Furthermore, replay of reward-related information in VS is enhanced during ripples [62]. Joint ensemble recordings showed that the HPC and VS replay their activity coherently, with hippocampal replay taking a leading role and VS reward-related information following in time [112]. Because joint replay occurs approximately ten times faster than the behavioral experience itself [112,140], this mechanism probably promotes spike-timing dependent synaptic plasticity (STDP) [141], which operates roughly in time windows 100–150 ms wide, and may mediate strengthening of place-reward associations during sleep-dependent memory consolidation. Independent pharmacological evidence has generally supported the role of VS in off-line processes subserving spatial memory consolidation [37].

its role in learning, or to its modulatory effects on performance, we note that even incentive-motivational influences may have to be acquired [70].

RL and the basal ganglia

How is it that the VS comes to generate signals predicting outcome properties, and how does hippocampal input contribute to this? To formulate a plausible model, we first consider RL as a general computational scheme for reward-dependent learning. The idea behind this class of algorithms is that a neural model (e.g. a connectionist network) processes sensory inputs and generates outputs that act on the environment, which subsequently feeds back a reinforcing signal to the model [71,72]. This signal is instrumental in adjusting the model's internal parameters (e.g. synaptic weights) to optimize its output with respect to a predefined goal, such as the maximization or prediction of reward [73].

A powerful class of RL algorithms is Temporal Difference (TD) learning [72]. A well-known variant of TD learning divides the computational tasks between a Critic and an Actor [74]. The Critic computes a reward–prediction signal, also known as a Value function (V_t) (Figure 4a). This is a single temporally continuous signal that fluctuates over time, depending on varying environmental cues or actions that inform the system to adjust its reward predictions as an animal pursues its goals. The reward–prediction signal is used to calculate an error in reward prediction, which (in simplified form) is done by subtracting the predicted outcome from the actual reward, once this is obtained. This error signal is used to improve the Critic's predictive performance, but can also instruct the Actor to optimize motor

Box 3. Outstanding questions

- Recent studies have shown the HPC–shell pathway to be critical for contextual conditioning and the BLA–core pathway for cue conditioning, whereas the hippocampal formation also projects to rostral parts of core and shell. What are the functions of the hippocampal–core pathway?
- VS neurons exhibit a great diversity of firing patterns, including responses to cues and reinforcers as well as correlates of motor behavior. Which of these patterns support outcome predictions and which ones do not?
- Although hippocampal firing activity tightly correlates with VS activity during active behavior and SWS, we do not yet understand how the HPC causally influences VS information processing. What is the effect of hippocampal lesions or inactivations on VS firing patterns?
- Glutamatergic afferents from neocortex, BLA and PFC that synapse on VS projection neurons exhibit long-term synaptic plasticity, but under what physiological conditions are LTP or LTP in these pathways induced *in vivo*? How does dopamine modulate this transmission and plasticity to induce regionally selective effects on behavior?

- Although current evidence supports the transfer of contextual information from the HPC to striatum, which other attributes of spatial–episodic information are conveyed? Does this transfer also include cue or object information, as well as temporal aspects of episodic memory and expected outcome?
- Which brain structures emit error signals to the striatum? Whereas error signaling by dopaminergic fibers is considered plausible, more work is needed to assess the functions of error coding in prefrontal structures, and its possible transfer to striatal target regions.
- How do different corticostriatal loops involving VS, DMS and DLS communicate under varying cognitive and behavioral demands? Much attention has been given to mesencephalic DA neurons as intermediate way stations from VS to DS (Figure 4b), but crosstalk may also take place via intrastriatal projections or at the cortical, thalamic and pallidal stages of information processing in loops.
- What is the precise definition of the VS, including its subregions, in the human brain, and what are the precise homologies between human and rodent striatum? Answering these questions will help to translate results from rodent work on VS to human psychopathology.
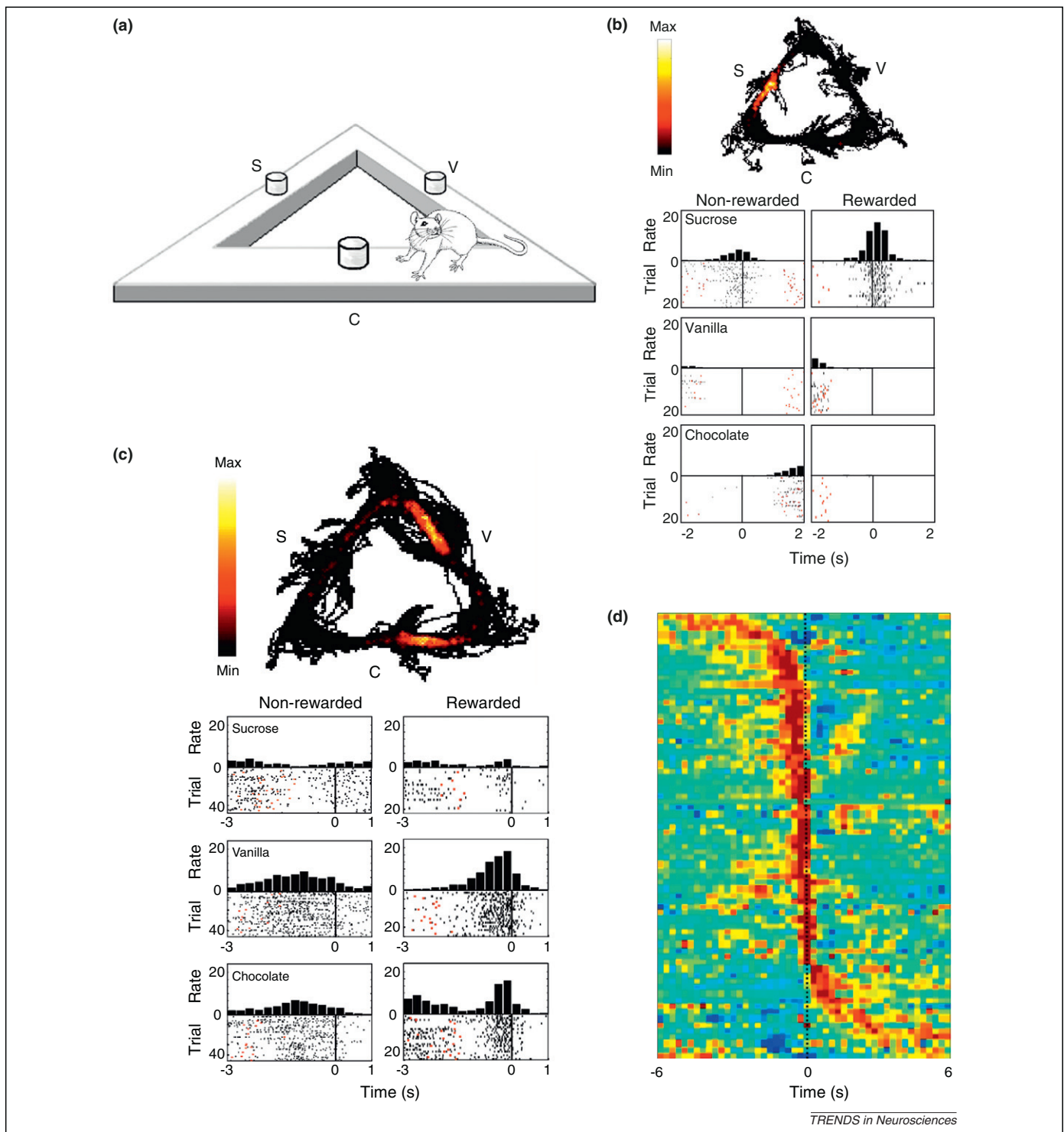
Figure 3. Firing patterns of rat ventral striatal (VS) neurons during foraging on a triangular track. **(a)** Behavioral task. Rats learned to run in a clockwise direction along a triangular track and encountered qualitatively different food rewards delivered to cups at fixed locations on each side of the triangle. The average probability of encountering a reward in a given cup was 0.33 (i.e. 1/3) and the rewards were distributed over time such that only one of the three cups was rewarded per lap. Meanwhile, ensemble recordings were made from the VS. Abbreviations: S, sucrose solution, delivered to cup at left side; V, vanilla pudding to cup at right side; C, chocolate mousse to cup at front side. **(b)** Single-unit firing pattern in the VS displayed a distinct firing response associated with only one reward site (i.e. S) [62]. Upper panel: rate map plotting firing rate as a function of the rat's position on the track. Firing rate is color-coded with highest rates (19 Hz at maximum) in white-yellow colors. Firing is virtually absent on most parts of the track. Lower panel: peri-event time histograms of the same neuron, synchronized on the rat's arrival at the three reward sites. Non-rewarded visits (left) are contrasted to rewarded visits (right) for each site. Black and red dots represent single spikes and arrivals at other reward sites, respectively, and are plotted as a function of trial number. Upper part of each subpanel denotes firing rate averaged over trials (in Hz). A ramp in firing rate is observed both in non-rewarded and rewarded trials, while the firing rate additionally increases just before arrivals at sucrose reward. **(c)** Different single-unit recording from the VS [62]. Plotting conventions are the same as in B. This cell generates ramps during approach to two of the three reward sites (i.e. V and C). Firing rate additionally increases shortly before reward receipt (i.e. time = 0 s) but rapidly drops after it. **(d)** Composition of 75 VS cells with task-related firing patterns from a different analysis of the same set of experiments in rats [119]. Only putative medium-sized spiny (i.e. projection) neurons are shown, whereas fast-spiking interneurons exhibit a different firing pattern. Z-Scored firing rates are color-coded and the 75 neurons are ordered from top to bottom according to the time of peak firing relative to reward site arrivals (at t = 0 s). A tessellation of all task phases is observed, with a concentration of peak rates shortly in advance of and at reward sites. Reproduced, with permission, from [62] (b and c) and [119] (d).
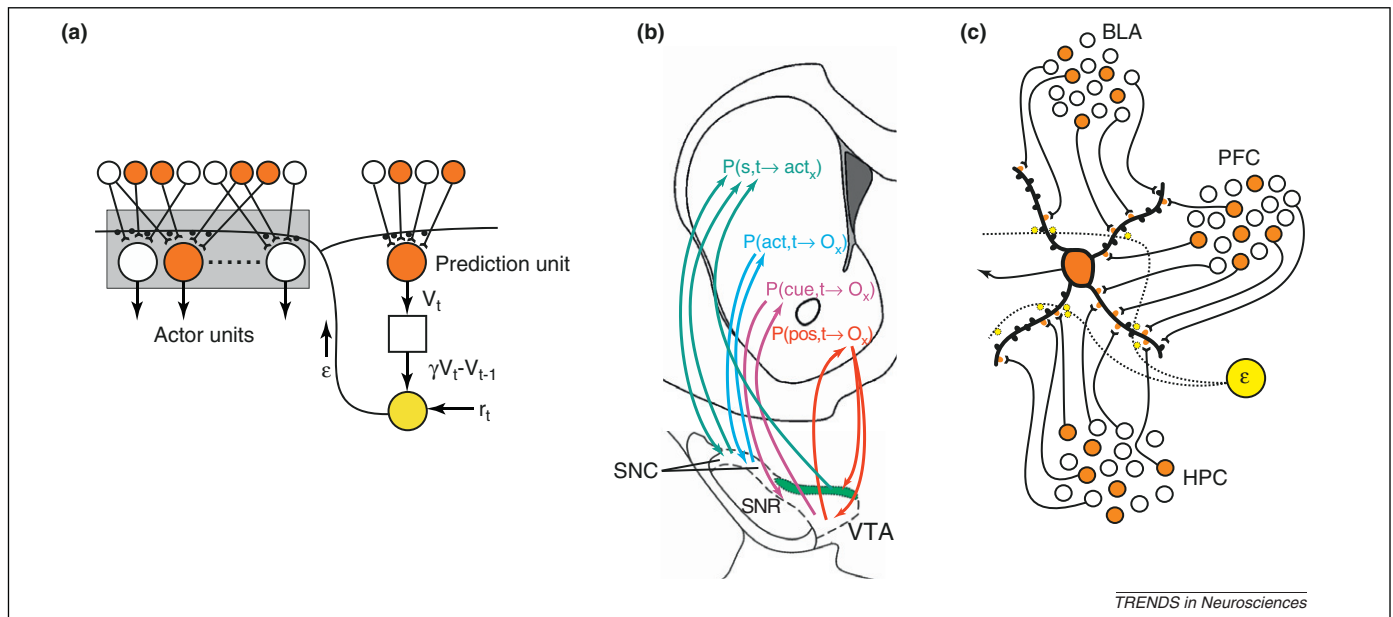
Figure 4. Classic Actor–Critic model and updated scheme for predictive learning in the striatum. **(a)** Actor–Critic network model consisting of a Prediction unit (Critic; right) and Actor units (left), both receiving inputs from cells in an afferent layer (active cells indicated in orange). Based on the sensory inputs the prediction unit receives via modifiable synapses, it emits a value signal (V_t) representing a time-varying prediction of summed future reward. Following computation of the temporal derivative of this signal ($\gamma V_t - V_{t-1}$, where γ is to discount the value of reward further ahead in the future), the lowermost unit (yellow) computes an error in reward prediction (ε) by summing up ($\gamma V_t - V_{t-1}$) with the actual reward r_t at time t. The error signal ε is broadcast to modifiable synapses connecting the input layer to Actor and Prediction units. Changes in synaptic strength are determined by activity of the error unit and a slowly decaying activity trace in input synapses. Adapted, with permission, from [74]. **(b)** Different types of predictive learning associated with striatal sectors in rat brain. The most ventromedial sector (shell; red) is predominated by time-varying outcome (O_x) prediction (P) based on position (pos) or context (subscript x indexes outcomes of a specific quality). The ventral striatum (VS) core (purple) generates outcome predictions based on discrete cues, whereas the hallmark of dorsomedial striatum (DMS; blue) is action–outcome learning. Stimulus–response learning in dorsolateral striatum (DLS; green) generates predictions about actions based on somatosensory (s) and motor information representing the organism's current postural and movement state. Thus, outcomes in DLS are specified as actions (act_x) of particular magnitude, speed and direction. The connectivity between striatal sectors and dopaminergic cell groups is predominantly reciprocal, but is supplemented with a projection from dorsal VTA and dorsomedial substantia nigra pars compacta (SNC) to DLS. No striatal projections to substantia nigra pars reticulata (SNR) are shown. Sections based on [118]. **(c)** Lesioning evidence indicates a predominance of different types of learning as in (b), but in addition, electrophysiological findings reveal convergence of inputs from different afferent sources on single neurons. Scheme depicts a medium-sized spiny neuron in VS receiving basolateral amygdala (BLA), prefrontal cortical (PFC) and hippocampal inputs, supplemented with prediction error (ε) information (dotted line, yellow unit). Inputs originate from ensembles of neurons activated (orange) by a specific cue (BLA), place (HPC) or task set (PFC). Synapses are modified when activated pre- and post-synaptically (orange) and when reached by the error signal (yellow). Error signals may be provided by dopaminergic neurons from the VTA [see (b)] or by glutamatergic sources such as agranular insular cortex (Figure 1).

output. Once it reaches its target areas, the prediction error functions as a teaching signal, influencing synaptic modifications in the Critic and Actor. Because a reward–prediction signal informs the organism what value or utility to expect based on its current state and actions, it can simultaneously serve as a measure of motivation to invigorate or attenuate actions. In the wider field of RL, TD learning exemplifies a model-free approach [75].

Within the brain, various structures qualify as candidates issuing positive or negatively reinforcing signals, including the amygdaloid complex, orbitofrontal cortex, striatum, habenula and mesencephalic dopaminergic neurons [76–81]. In a broad sense, a multiplicity of structures and plasticity mechanisms are probably involved in RL, some of which hinge on glutamatergic transmission [82,83] and others on dopamine or other neuromodulatory systems [84]. Nonetheless, the analogy between firing patterns of dopamine neurons and the error-coding module in TD learning is particularly striking, supporting the previously proposed hypothesis that the firing rate of dopamine neurons signals a reward prediction error [85].

It is less than clear, however, what the role of the striatum and afferent cortical areas might be in this scheme. Earlier models of Actor–Critic architectures proposed that neurons in the striosomes of DS [16] function as

Critic, whereas matrix neurons serve as Actors [76,86]. Alternatively, the VS might serve as Critic and the DS as Actor [76,87]. Accordingly, the VS would supply dopamine neurons with value signals, which subsequently broadcast error signals to the striatum to improve outcome predictions and stimulus–response learning. As predicted by TD learning, dopamine may regulate plasticity of corticostriatal synapses ([88–90], but see [91,92] for absence of dopamine effects in ventral striatal preparations).

Electrophysiological evidence [59,62,63,65,68,69] supports a role of VS in outcome prediction, although not necessarily as modeled by an RL Critic. Anatomically, VS outputs to VTA and SNC [13,93–95] support the idea that outcome–expectancy signals coded by VS projection neurons may modulate firing of dopaminergic neurons. However, orbitofrontal and other prefrontal inputs may also be important for generating error-like signals in dopamine neurons [77]. Conversely, the VS is supplied with a rich dopaminergic innervation from VTA and medial SNC [13,93–95], consistent with the idea that it receives error signals that may modulate or modify corticostriatal synapses. Despite these consistencies with a role of VS as Critic, current evidence suggests that the Actor–Critic model of VS and DS should be replaced by an alternative scheme.

VS representation of outcome predictions: a revised scheme

A number of findings suggest that an alternative scheme for explaining computational functions of the VS may be warranted. First, the electrophysiological evidence discussed above suggests that outcome-predictive activity in VS is not a monolithic, uniform signal. Although firing-rate ramps exemplify how a cue-dependent value function may be neurally expressed, the more common behavioral situation is that the outcome is preceded by multiple cues and actions set in a specific context. Here, there is no fixed set of VS neurons that continuously expresses a single value signal over time, but instead sequentially activated neuronal ensembles are found, encoding successive task elements (Figure 3d and Box 2), compatible with DS single-unit recordings [60]. These findings are consistent with the use of ensemble coding [62,96,97] for signaling outcome predictions, temporally chained as in a relay race, each coding for valuable task elements leading to the outcome. Nonetheless, such configurations are still compatible with a scheme in which error signals can be computed in areas downstream to the VS, such as the VTA and SNC.

Second, current evidence is incongruent with a strict segregation of tasks between VS and DS as in an Actor–Critic architecture. Despite major connectional and functional differences, DS and VS share the same basic design of corticostriatal loops, with the entire striatum receiving topographic dopaminergic inputs but also feeding back output to the mesencephalic area of origin, in partially closed striato–nigro–striatal loops [93–95]. This fundamental resemblance between VS and DS suggests a corresponding similarity in computational function, whereas the inputs (or informational contents) used in the computation are different. An Actor–Critic division suggests that learning in the Actor necessarily depends on learning by the Critic (Figure 4a), but important evidence argues that Action–Outcome learning mediated by the DMS can proceed in the presence of VS lesions [98]. The Actor–Critic scheme also assumes that the error is broadcast uniformly to both Actor and Critic (Figure 4a). Evidence in primates and rats suggests that VTA neurons receiving VS inputs project to DS regions, but this direct projection only reaches limited zones in the DS [93–95] (Figure 4b). The predominant pattern remains topographic: the DS is innervated by the lateral SNC, whereas the VS core and shell receive inputs from the VTA and medial SNC [13,80,93,94] (Figure 4b).

Our alternative scheme (Figure 4b,c) not only takes into account outcome-predictive coding in the VS, but also in the DMS and DLS [60,99,100]. As a consequence, it assumes that striatal neurons share a fundamental function in coding outcome predictions. However, the main VS–DMS–DLS differences lie in the sources used to compute predictions, and hence, in the informational domains to which predictions pertain. In this context, it is relevant that the DMS also receives hippocampal and amygdaloid input, albeit sparser than in VS, whereas the DLS is virtually devoid of these projections [13,19]. Returning to the question of how hippocampal inputs shape VS predictive activity, we first note that these inputs are mediated by efficacious, glutamatergic synapses [96] and probably convey information about spatial context to help sculpt temporal firing patterns of VS neurons. Secondly, converging with hippocampal input, the BLA provides information to VS about discrete stimuli and the medial PFC codes information about task rules and set, behavioral strategies and planning [101–104] (Figure 4c). Accordingly, VS outcome predictions will be primarily based on spatial context, discrete cues and task set. By contrast, the DLS receives information from the somatosensory cortices, primary and higher motor cortices, coding the preparation, execution and sensing of specific movements [9,13,105], whereas the DMS processes inputs from dorsolateral and medial PFC and anterior cingulate cortex pertaining to more global cognitive and motor operations [4,15,17]. Thus, outcome predictions in DLS and DMS are primarily derived from specific or global sensorimotor processing, as observed in single-unit firing during or in advance of movements [99,100]. This concept agrees with the implication of DMS, but not VS, in action–outcome learning. Its logical consequence is that the DLS, implicated in stimulus–response learning, uses detailed sensory and motor inputs to predict a non-motivational type of outcome, viz. a specific action (Figure 4b).

The notion of a common architecture accommodating different information domains can be extended to the ventral mesencephalon. Following earlier work suggesting that dopaminergic neurons not only transmit reward-related signals but also signal novelty, saliency and surprise, including information about aversive events [106,107], two types of dopamine neurons were recently discerned in the primate mesencephalon [80], one coding motivational value (showing opposite responses to appetitive versus aversive events) and the other coding motivational salience (responding similarly to appetitive and aversive events). Dopamine signaling in the ventromedial mesencephalon reaching the shell and ventromedial PFC was proposed to depend on signed value information (i.e. of opposite sign for positive and negative value). By contrast, the informational domain of dorsolateral dopamine cells, projecting to the core, DMS, DLS and dorsolateral PFC, comprises salient and surprising events in general.

This heterogeneity of dopamine neurons is associated with a multiplicity of cellular functions of dopamine in the striatum. Besides dopamine effects on long-term synaptic plasticity, many reversible effects have been described, often showing differences between DS and VS [108]. Focusing on VS, dopamine exerts reversible, suppressive control over both glutamatergic excitation and GABAergic inhibition, including communication between projection neurons [96,108,109]. Hippocampal and prefrontal inputs to VS are differentially controlled by D1 and D2-type dopamine receptors [110], suggesting how dopamine may gate different afferent sources, bias outcome predictions and invigorate different behaviors. How these mechanisms exactly modulate limbic control over motivated behavior is unclear (Box 3), but the framework of functionally distinct, potentially competing ensembles in shell and core, differentially innervated by glutamatergic sources and controlled by dopamine [96], has recently gained support [47,62,97].

In addition to temporally evolving value representations in VS ensembles, and the idea of outcome predictions based on different informational domains, a third deviation from Actor–Critic schemes lies in the conditional nature of outcome-predictive striatal signals and it is especially here that hippocampal inputs become crucial to consider. Results from the CPP experiments using lesioned animals [42] (see above) may be explained by the conversion of spatial–episodic hippocampal information into a VS signal coding for impending reward close to the animal's current position (Figure 4b,c). This transformation can be accomplished by adjusting the weights of hippocampal synapses onto VS projection neurons by long-term potentiation (LTP) and long-term depression (LTD)-like processes, as documented for related corticostriatal pathways [89,91,92] (Figure 4c and Box 2). If a location has been consistently paired with an unpredicted outcome, hippocampal–VS connections are proposed to be associatively strengthened, promoting firing of VS neurons given that context. This alone, however, does not generally suffice to explain many behavioral and electrophysiological observations. Task execution is not only set within a spatial context, but is usually initiated by a discrete cue and follows a specific layout of rules and contingencies. For example, in a task where an instrumental locomotor response is required to reach a cued goal, an approach response only makes sense if the rat is away from the goal site. Thus, place-specific HPC ensembles will be coactivated with cue-specific BLA and rule-specific medial PFC ensembles. This configuration is supported by studies showing convergent excitatory inputs to single VS units (e.g. [111]) (Figure 4c). Time-varying predictive signals in VS will thus be conditional on multiple inputs, converging and temporally summating to a variable extent across cell populations.

Fourthly, VS neurons appear to be not only sensitive to value (or utility), but also to the identity and spatial location of the outcome [62,66,67] (Figures 3 and 4b). The spatial–episodic HPC input, which plausibly contributes state-specific information to the VS, is an important factor in this respect [69,112]. Thus, outcome-predictive signals in VS (but also in DMS and DLS) are envisioned to be more multi-dimensional than the scalar value signal posited by classical RL models, both in terms of predictors and outcome specificity. In other words, VS signals not only predict how good or bad the outcome will be, but also the 'What' and 'Where' of it, and possibly further aspects such as when it will come. Such outcome-specific predictions may be coded in parallel with more general common-currency value signals.

Where, in this alternative scheme, could the Actor posited by RL models be situated? Several possibilities can be suggested, although it is difficult to assess their validity at present. The idea that the matrix of DS functions as Actor [76,86] is difficult to evaluate but deserves further testing. Secondly, the concept that striatal subregions compute their own predictions by striato–nigro–striatal loops (Figure 4b) is not incompatible with a coexisting influence from ventromedial towards more dorsolateral sensorimotor processing circuits [93,94,113]. Thirdly, Actor modules may be situated downstream from striatum but within the basal ganglia. Pallidal structures and the SNR are interesting candidates here because they are densely innervated both

by striatal projection neurons and dopaminergic fibers and have been implicated in action selection and execution [114,115]. Fourthly, the basal ganglia are not the exclusive domain for motor learning; actions may be acquired elsewhere in the brain, for instance in premotor–motor corticothalamic and cerebellar networks [116,117], whereas the striatum may then function to compute outcome predictions that affect action learning and action selection in these networks. Regardless of this debate, the current model holds that VS projection neurons code specific outcome-predicting signals which, at the same time, act to invigorate or disinhibit particular motor patterns by their effects on downstream areas, such as SNR, lateral hypothalamus, brain stem and ventral pallidal targets and connected thalamo-cortical feedback loops.

Concluding remarks

Should the VS, in the end, be classified as a component of the episodic or procedural memory system? The most parsimonious answer holds that it incorporates components of both types of memory, constituting a third system integrating inputs from the BLA, HPC, PFC and other areas to generate motivational (outcome-predictive) signals that act on downstream motor systems to invigorate or disinhibit goal-directed behaviors. Compared with DLS and DMS, the VS is distinguished by its role in how discrete cues and contexts come to exert pavlovian control over behaviors during learning. Current data imply the HPC–shell pathway in contextual conditioning and spatial processing, and the BLA–core pathway in cue conditioning [42], although the complexity of the system suggests additional as yet unknown functions and leaves open the possibility of functional overlap [43,44] (Box 3).

We propose that recent experimental data are in agreement with a need to replace the classical Actor–Critic RL with a revised scheme. The key elements in this scheme hold that: (i) outcome-predictive activity in VS (and probably DS) is expressed by sequentially activated ensembles; (ii) VS, DMS and DLS operate to generate outcome predictions according to the same principles but in different informational domains, working in parallel but also interacting in a ventromedial to dorsolateral direction; (iii) outcome-predictive signaling in the striatum is of a conditional and combinatorial nature, as illustrated by the convergence of spatial, cue- and rule-specific information on VS ensembles during contextual and cue conditioning tasks; and (iv) based on the multi-dimensional nature of information carried by its inputs, VS coding incorporates episodic features of the outcome, such as reward quality or location, in addition to scalar value representations. Generation of outcome predictions is proposed to rely on synaptic plasticity mechanisms boosted during slow-wave sleep (SWS) (Box 2). Altogether, these features place our scheme closer to model-based architectures for RL than previously envisioned in model-free approaches [75].

References

1 Milner, B. *et al.* (1998) Cognitive neuroscience and the study of memory. *Neuron* 20, 445–468

2 Packard, M.G. and McGaugh, J.L. (1992) Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: further evidence for multiple memory systems. *Behav. Neurosci.* 106, 439–446

3 McDonald, R.J. and White, N.M. (1993) A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav. Neurosci.* 107, 3–22

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

5 Tulving, E. and Markowitsch, H.J. (1998) Episodic and declarative memory: role of the hippocampus. *Hippocampus* 8, 198–204

6 Eichenbaum, H. *et al.* (2007) The medial temporal lobe and recognition memory. *Ann. Rev. Neurosci.* 30, 123–152

7 Wang, S.H. and Morris, R.G. (2010) Hippocampal–neocortical interactions in memory formation, consolidation, and reconsolidation. *Ann. Rev. Psychol.* 61, 49–79

8 Squire, L.R. and Wixted, J.T. (2011) The cognitive neuroscience of human memory since H.M. *Ann. Rev. Neurosci.* 34, 259–288

9 Alexander, G.E. *et al.* (1990) Basal ganglia–thalamocortical circuits: parallel substrates for motor, oculomotor, 'prefrontal' and 'limbic' functions. *Prog. Brain Res.* 85, 119–146

10 Nambu, A. (2008) Seven problems on the basal ganglia. *Curr. Opin. Neurobiol.* 18, 595–604

11 Kopell, B.H. and Greenberg, B.D. (2008) Anatomy and physiology of the basal ganglia: implications for DBS in psychiatry. *Neurosci. Biobehav. Rev.* 32, 408–422

12 Corbit, L.H. and Balleine, B.W. (2003) The role of prelimbic cortex in instrumental conditioning. *Behav. Brain Res.* 146, 145–157

13 Voorn, P. *et al.* (2004) Putting a spin on the dorsal–ventral divide of the striatum. *Trends Neurosci.* 27, 468–474

14 Everitt, B.J. and Robbins, T.W. (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489

15 Yin, H.H. *et al.* (2005) The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523

16 Graybiel, A.M. (2008) Habits, rituals, and the evaluative brain. *Ann. Rev. Neurosci.* 31, 359–387

17 Balleine, B.W. *et al.* (2009) The integrative function of the basal ganglia in instrumental conditioning. *Behav. Brain Res.* 199, 43–52

18 Heimer, L. *et al.* (1985) Basal Ganglia. In *The Rat Nervous System* (Vol. 1) (Paxinos, A., ed.), pp. 37–86, Academic Press

19 Groenewegen, H.J. *et al.* (1987) Organization of the projections from the subiculum to the ventral striatum in the rat. A study using anterograde transport of *Phaseolus vulgaris* leucoagglutinin. *Neuroscience* 23, 103–120

20 Moser, E. *et al.* (1993) Spatial learning impairment parallels the magnitude of dorsal hippocampal lesions, but is hardly present following ventral lesions. *J. Neurosci.* 13, 3916–3925

21 Bannerman, D.M. *et al.* (2004) Regional dissociations within the hippocampus: memory and anxiety. *Neurosci. Biobehav. Rev.* 28, 273–283

22 Good, M. and Honey, R.C. (1991) Conditioning and contextual retrieval in hippocampal rats. *Behav. Neurosci.* 105, 499–509

23 Fanselow, M.S. and Poulos, A.M. (2005) The neuroscience of mammalian associative learning. *Ann. Rev. Psychol.* 56, 207–234

24 Phillips, R.G. and LeDoux, J.E. (1992) Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav. Neurosci.* 106, 274–285

25 Ito, R. *et al.* (2006) Selective excitotoxic lesions of the hippocampus and basolateral amygdala have dissociable effects on appetitive cue and place conditioning based on path integration in a novel Y-maze procedure. *Eur. J. Neurosci.* 23, 3071–3080

26 Quinn, J.J. *et al.* (2008) Dorsal hippocampus involvement in delay fear conditioning depends upon the strength of the tone–footshock association. *Hippocampus* 18, 640–654

27 Levita, L. and Muzzio, I.A. (2010) Role of the hippocampus in goal-oriented tasks requiring retrieval of spatial versus non-spatial information. *Neurobiol. Learn. Mem.* 93, 581–588

28 Moser, M.B. and Moser, E.I. (1998) Functional differentiation in the hippocampus. *Hippocampus* 8, 608–619

29 Royer, S. *et al.* (2010) Distinct representations and theta dynamics in dorsal and ventral hippocampus. *J. Neurosci.* 30, 1777–1787

30 Bast, T. *et al.* (2003) Dorsal hippocampus and classical fear conditioning to tone and context in rats: effects of local NMDA-receptor blockade and stimulation. *Hippocampus* 13, 657–675

31 Kjelstrup, K.B. *et al.* (2008) Finite scale of spatial representation in the hippocampus. *Science* 321, 140–143

32 Bast, T. (2011) The hippocampal learning–behavior translation and the functional significance of hippocampal dysfunction in schizophrenia. *Curr. Opin. Neurobiol.* 21, 492–501

33 Mogenson, G.J. *et al.* (1980) From motivation to action: functional interface between the limbic system and the motor system. *Prog. Neurobiol.* 14, 69–97

34 Annett, L.E. *et al.* (1989) The effects of ibotenic acid lesions of the nucleus accumbens on spatial learning and extinction in the rat. *Behav. Brain Res.* 31, 231–242

35 Seamans, J.K. and Phillips, A.G. (1994) Selective memory impairments produced by transient lidocaine-induced lesions of the nucleus accumbens in rats. *Behav. Neurosci.* 108, 456–468

36 Floresco, S.B. *et al.* (1997) Selective roles for hippocampal, prefrontal cortical, and ventral striatal circuits in radial-arm maze tasks with or without a delay. *J. Neurosci.* 17, 1880–1890

37 Ferretti, V. *et al.* (2010) Ventral striatal plasticity and spatial memory. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7945–7950

38 McNally, G.P. and Westbrook, R.F. (2006) Predicting danger: the nature, consequences, and neural mechanisms of predictive fear learning. *Learn. Mem.* 13, 245–253

39 Robbins, T.W. *et al.* (2008) Drug addiction and the memory systems of the brain. *Ann. N. Y. Acad. Sci.* 1141, 1–21

40 Everitt, B.J. *et al.* (1991) The basolateral amygdala–ventral striatal system and conditioned place preference: further evidence of limbic–striatal interactions underlying reward-related processes. *Neuroscience* 42, 1–18

41 Pezze, M.A. *et al.* (2002) Increased conditioned fear response and altered balance of dopamine in the shell and core of the nucleus accumbens during amphetamine withdrawal. *Neuropharmacology* 42, 633–643

42 Ito, R. *et al.* (2008) Functional interaction of the hippocampus and nucleus accumbens shell is necessary for the acquisition of appetitive spatial context conditioning. *J. Neurosci.* 28, 6950–6959

43 Levita, L. *et al.* (2002) Disruption of pavlovian contextual conditioning by excitotoxic lesions of the nucleus accumbens core. *Behav. Neurosci.* 116, 539–552

44 Maldonado-Irizarry, C.S. and Kelley, A.E. (1995) Excitatory amino acid receptors within nucleus accumbens subregions differentially mediate spatial learning in the rat. *Behav. Pharmacol.* 6, 527–539

45 Di Ciano, P. and Everitt, B.J. (2004) Direct interactions between the basolateral amygdala and nucleus accumbens core underlie cocaine-seeking behavior by rats. *J. Neurosci.* 24, 7167–7173

46 Taylor, J.R. and Robbins, T.W. (1984) Enhanced behavioral control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. *Psychopharmacology (Berl.)* 84, 405–412

47 Ito, R. and Hayen, A. (2011) Opposing roles of nucleus accumbens core and shell dopamine in the modulation of limbic information processing. *J. Neurosci.* 31, 6001–6007

48 O'Keefe, J. and Nadel, L. (1978) *The Hippocampus as a Cognitive Map*, Oxford University Press

49 de Hoz, L. and Wood, E.R. (2006) Dissociating the past from the present in the activity of place cells. *Hippocampus* 16, 704–715

50 Leutgeb, S. *et al.* (2005) Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science* 309, 619–623

51 Ergorul, C. and Eichenbaum, H. (2006) Essential role of the hippocampal formation in rapid learning of higher-order sequential associations. *J. Neurosci.* 26, 4111–4117

52 Groenewegen, H.J. *et al.* (1982) Cortical afferents of the nucleus accumbens in the cat, studied with anterograde and retrograde transport techniques. *Neuroscience* 7, 977–996

53 Witter, M.P. and Groenewegen, H.J. (1986) Connections of the parahippocampal cortex in the cat. IV. Subcortical efferents. *J. Comp. Neurol.* 252, 51–77

54 Sharp, P.E. (2006) Subicular place cells generate the same 'map' for different environments: comparison with hippocampal cells. *Behav. Brain Res.* 174, 206–214

55 Lever, C. *et al.* (2009) Boundary vector cells in the subiculum of the hippocampal formation. *J. Neurosci.* 29, 9771–9777

56 Derdikman, D. and Moser, E.I. (2010) A manifold of spatial maps in the brain. *Trends Cogn. Sci.* 14, 561–569

57 Murray, E.A. *et al.* (2007) Visual perception and memory: a new view of medial temporal lobe function in primates and rodents. *Ann. Rev. Neurosci.* 30, 99–122

58 Shidara, M. *et al.* (1998) Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials. *J. Neurosci.* 18, 2613–2625

59 Setlow, B. *et al.* (2003) Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* 38, 625–636

60 Schultz, W. (2006) Behavioral theories and the neurophysiology of reward. *Ann. Rev. Psychol.* 57, 87–115

61 Taha, S.A. *et al.* (2007) Cue-evoked encoding of movement planning and execution in the rat nucleus accumbens. *J. Physiol.* 584, 801–818

62 Lansink, C.S. *et al.* (2008) Preferential reactivation of motivationally valuable information in the ventral striatum. *J. Neurosci.* 28, 6372–6382

63 Roesch, M.R. *et al.* (2009) Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J. Neurosci.* 29, 13365–13376

64 Jensen, J. *et al.* (2003) Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 40, 1251–1257

65 Roitman, M.F. *et al.* (2005) Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* 45, 587–597

66 Wheeler, R.A. and Carelli, R.M. (2009) Dissecting motivational circuitry to understand substance abuse. *Neuropharmacology* 56, 149–159

67 McDannald, M.A. *et al.* (2011) Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J. Neurosci.* 31, 2700–2705

68 Schultz, W. *et al.* (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. *J. Neurosci.* 12, 4595–4610

69 Van der Meer, M.A. and Redish, A.D. (2011) Theta phase precession in rat ventral striatum links place and reward information. *J. Neurosci.* 31, 2843–2854

70 Robbins, T.W. and Everitt, B.J. (2007) A role for mesencephalic dopamine in activation: commentary on Berridge. *Psychopharmacology* 191, 433–437

71 Widrow, B. *et al.* (1973) Punish/reward: learning with a critic in adaptive threshold systems. *IEEE Transact. Syst. Man Cybern.* 3, 455–465

72 Sutton, R.S. and Barto, A.G. (1998) *Reinforcement Learning*, MIT Press

73 Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning. II: Current Research and Theory* (Black, A.H. and Prokasy, W.F., eds), pp. 64–99, Appleton Century Crofts

74 Barto, A.G. (1995) Adaptive critics and the basal ganglia. In *Models of Information Processing in the Basal Ganglia* (Houk, J.C. *et al.*, eds), pp. 215–232, MIT Press

75 Daw, N.D. *et al.* (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711

76 Schultz, W. (1998) Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27

77 Takahashi, Y.K. *et al.* (2009) The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* 62, 269–280

78 Van Duuren, E. *et al.* (2009) Single-cell and population coding of expected reward probability in rat orbitofrontal cortex. *J. Neurosci.* 29, 8965–8976

79 Sul, J.H. *et al.* (2010) Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* 66, 449–460

80 Bromberg-Martin, E.S. *et al.* (2010) Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834

81 Morrison, S.E. and Salzman, C.D. (2010) Re-valuing the amygdala. *Curr. Opin. Neurobiol.* 20, 221–230

82 Pennartz, C.M.A. (1997) Reinforcement learning by Hebbian synapses with adaptive thresholds. *Neuroscience* 81, 303–319

83 Kelley, A.E. *et al.* (2003) Glutamate-mediated plasticity in corticostriatal networks: role in adaptive motor learning. *Ann. N. Y. Acad. Sci.* 1003, 159–168

84 Yu, A.J. and Dayan, P. (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692

85 Schultz, W. (2007) Multiple dopamine functions at different time courses. *Ann. Rev. Neurosci.* 30, 259–288

86 Houk, J.C. *et al.* (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Models of Information Processing in the Basal Ganglia* (Houk, J.C. *et al.*, eds), pp. 249–270, MIT Press

87 Attalah, H.E. *et al.* (2007) Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat. Neurosci.* 10, 126–131

88 Arbuthnott, G.W. and Wickens, J. (2007) Space, time and dopamine. *Trends Neurosci.* 30, 62–69

89 Mahon, S. *et al.* (2004) Corticostriatal plasticity: life after the depression. *Trends Neurosci.* 27, 460–467

90 Pawlak, V. and Kerr, J.N. (2008) Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neurosci.* 28, 2435–2446

91 Pennartz, C.M.A. *et al.* (1993) Synaptic plasticity in an in vitro slice preparation of the rat nucleus accumbens. *Eur. J. Neurosci.* 5, 107–117

92 Kauer, J.A. and Malenka, R.C. (2007) Synaptic plasticity and addiction. *Nat. Rev. Neurosci.* 8, 844–858

93 Haber, S.N. and Knutson, B. (2010) The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26

94 Maurin, Y. *et al.* (1999) Three-dimensional distribution of nigrostriatal neurons in the rat: relation to the topography of striatonigral projections. *Neuroscience* 91, 891–909

95 Pennartz, C.M.A. *et al.* (2009) Corticostriatal interactions during learning, memory processing and decision-making. *J. Neurosci.* 29, 12831–12838

96 Pennartz, C.M.A. *et al.* (1994) The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomic data. *Progr. Neurobiol.* 42, 719–761

97 Koya, E. *et al.* (2009) Targeted disruption of cocaine-activated nucleus accumbens neurons prevents context-specific sensitization. *Nat. Neurosci.* 12, 1069–1073

98 de Borchgrave, R. *et al.* (2002) Effects of cytotoxic nucleus accumbens lesions on instrumental conditioning in rats. *Exp. Brain Res.* 144, 50–68

99 Hikosaka, O. *et al.* (2008) New insights on the subcortical representation of reward. *Curr. Opin. Neurobiol.* 18, 203–208

100 Samejima, K. *et al.* (2005) Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340

101 Duff, A. *et al.* (2011) A biologically based model for the integration of sensory-motor contingencies in rules and plans: a prefrontal cortex based extension of the Distributed Adaptive Control architecture. *Brain Res. Bull.* 85, 289–304

102 Muhammad, R. *et al.* (2006) A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *J. Cogn. Neurosci.* 18, 974–989

103 Rich, E.L. and Shapiro, M. (2009) Rat prefrontal cortical neurons selectively code strategy switches. *J. Neurosci.* 29, 7208–7219

104 Durstewitz, D. *et al.* (2010) Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* 66, 438–448

105 Pidoux, M. *et al.* (2011) Integration and propagation of somatosensory responses in the corticostriatal pathway: an intracellular study in vivo. *J. Physiol.* 589, 263–281

106 Ljungberg, T. *et al.* (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163

107 Redgrave, P. *et al.* (2008) What is reinforced by phasic dopamine signals? *Brain Res. Rev.* 58, 322–339

108 Moyer, J.T. *et al.* (2007) Effects of dopaminergic modulation on the integrative properties of the ventral striatal medium spiny neuron. *J. Neurophysiol.* 98, 3731–3748

109 Taverna, S. *et al.* (2005) Dopamine D1-receptors modulate lateral inhibition between principal cells of the nucleus accumbens. *J. Neurophysiol.* 93, 1816–1819

110 Goto, Y. and Grace, A.A. (2008) Limbic and cortical information processing in the nucleus accumbens. *Trends Neurosci.* 31, 552–558

111 Floresco, S.B. *et al.* (2001) Modulation of hippocampal and amygdalar-evoked activity of nucleus accumbens neurons by dopamine: cellular mechanisms of input selection. *J. Neurosci.* 21, 2851–2860

112 Lansink, C.S. *et al.* (2009) Hippocampus leads ventral striatum in replay of place–reward information. *PLoS Biol.* 7, e1000173

113 Belin, D. and Everitt, B.J. (2008) Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron* 57, 432–441

114 Deniau, J.M. *et al.* (2007) The pars reticulata of the substantia nigra: a window to basal ganglia output. *Prog. Brain Res.* 160, 151–172

115 Rommelfanger, K.S. and Wichmann, T. (2010) Extrastriatal dopaminergic circuits of the basal ganglia. *Front. Neuroanat.* 4, 139

116 Arce, F. *et al.* (2010) Combined adaptiveness of specific motor cortical ensembles underlies learning. *J. Neurosci.* 30, 5415–5425

117 Dean, P. *et al.* (2010) The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nat. Rev. Neurosci.* 11, 30–43

118 Paxinos, G. and Watson, C., eds (2007) *The Rat Brain in Stereotaxic Coordinates* (6th edn), Academic Press

119 Lansink, C.S. *et al.* (2010) Fast spiking interneurons of the rat ventral striatum: temporal coordination of activity with principal cells and responsiveness to reward. *Eur. J. Neurosci.* 32, 494–508

120 Russo, S.J. *et al.* (2010) The addicted synapse: mechanisms of synaptic and structural plasticity in nucleus accumbens. *Trends Neurosci.* 33, 267–276

121 Breiter, H.C. *et al.* (1997) Acute effects of cocaine on brain activity and emotion. *Neuron* 19, 591–611

122 Volkow, N.D. *et al.* (2011) Addiction: beyond dopamine reward circuitry. *Proc. Natl. Acad. Sci. U.S.A.* DOI: 10.1073/pnas.1010654108

123 Beck, A. *et al.* (2009) Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. *Biol. Psychiatry* 66, 734–742

124 Reuter, J. *et al.* (2005) Pathological gambling is linked to reduced activation of the mesolimbic reward system. *Nat. Neurosci.* 8, 147–148

125 Scheres, A. *et al.* (2007) Ventral striatal hyporesponsiveness during reward anticipation in attention deficit/hyperactivity disorder. *Biol. Psychiatry* 61, 720–724

126 Knutson, B. *et al.* (2008) Neural responses to monetary incentives in major depression. *Biol. Psychiatry* 63, 686–692

127 Dagher, A. and Robbins, T.W. (2009) Personality, addiction, dopamine: insights from Parkinson's disease. *Neuron* 61, 502–510

128 Kapur, S. (2003) Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* 160, 13–23

129 Grace, A.A. (2010) Ventral hippocampus, interneurons and schizophrenia: a new understanding of the pathophysiology of schizophrenia and its implications for treatment and prevention. *Curr. Dir. Psychol. Sci.* 19, 232–237

130 Kegeles, L.S. *et al.* (2010) Increased synaptic dopamine in associative regions of the striatum in schizophrenia. *Arch. Gen. Psychiatry* 67, 231–239

131 Berke, J.D. *et al.* (2004) Oscillatory entrainment of striatal neurons in freely moving rats. *Neuron* 43, 883–896

132 O'Keefe, J. and Burgess, N. (2005) Dual phase and rate coding in hippocampal place cells: theoretical significance and relationship to entorhinal grid cells. *Hippocampus* 15, 853–866

133 Maurer, A.P. and McNaughton, B.L. (2007) Network and intrinsic cellular mechanisms underlying theta phase precession of hippocampal neurons. *Trends Neurosci.* 30, 325–333

134 Johnson, A. and Redish, A.D. (2007) Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* 27, 12176–12189

135 Van der Meer, M.A. and Redish, A.D. (2009) Covert expectation-of-reward in rat ventral striatum at decision points. *Front. Integr. Neurosci.* 3, 1

136 Sullivan, D. *et al.* (2011) Relationships between hippocampal sharp waves, ripples, and fast gamma oscillation: influence of dentate and entorhinal cortical activity. *J. Neurosci.* 31, 8605–8616

137 Wilson, M.A. and McNaughton, B.L. (1994) Reactivation of hippocampal ensemble memories during sleep. *Science* 265, 676–679

138 Girardeau, G. *et al.* (2009) Selective suppression of hippocampal ripples impairs spatial memory. *Nat. Neurosci.* 12, 1222–1223

139 Pennartz, C.M.A. *et al.* (2004) The ventral striatum in off-line processing: ensemble reactivation during sleep and modulation by hippocampal ripples. *J. Neurosci.* 24, 6446–6456

140 Euston, D.R. *et al.* (2007) Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science* 318, 1147–1150

141 Caporale, N. and Dan, Y. (2008) Spike timing-dependent plasticity: a Hebbian learning rule. *Ann. Rev. Neurosci.* 31, 25–46