Vicarious trial and error

A. David Redish

Abstract | When rats come to a decision point, they sometimes pause and look back and forth as if deliberating over the choice; at other times, they proceed as if they have already made their decision. In the 1930s, this pause-and-look behaviour was termed 'vicarious trial and error' (VTE), with the implication that the rat was 'thinking about the future'. The discovery in 2007 that the firing of hippocampal place cells gives rise to alternating representations of each of the potential path options in a serial manner during VTE suggested a possible neural mechanism that could underlie the representations of future outcomes. More-recent experiments examining VTE in rats suggest that there are direct parallels to human processes of deliberative decision making, working memory and mental time travel.

Neural ensembles
A set of cells recorded separably, but simultaneously, generally from a single brain structure during behaviour. Information represented within ensembles can be decoded from a sufficiently large ensemble.

Mental time travel
A process in which one imagines another time and place. Sometimes referred to as episodic future thinking.

Schema
An expertise-dependent representation of the structure of the world, identifying the important parameters over which the world varies.

Cognitive map
A world representation on which one can plan. Currently, the term is generally used in a spatial context, but Tolman's original use of the term was closer to the current use of the word 'schema'.

Department of Neuroscience, University of Minnesota, Minneapolis, Minnesota 55455, USA.
redish@umn.edu

doi:10.1038/nrn.2015.30
Published online 19 Feb 2016

In the 1930s, researchers such as Muenzinger, Gentry and Tolman noticed that, at the choice point of a maze, rats would occasionally pause and look back and forth as though confused about which way to go[1–3]. These researchers speculated that rats were imagining potential future options, and called this behaviour 'vicarious trial and error' (VTE)[1–3]. The researchers' hypotheses that the rats were mentally searching future trajectories contrasted with simpler situation–action-chain hypotheses that maintained that an animal recognized a situation and 'released' a well-learned series of actions[4] (FIG. 1).

In the 1940s, search hypotheses as an explanation for behaviour lost favour to situation–action-chain hypotheses, in part because the proponents of the former could not provide mechanistic explanations for how VTE behaviour related to decisions. Concepts of computation[5,6], information[7] and representation[5] were only just being developed in the 1940s. Moreover, the computer revolution that found that the extraction of information from representations (for example, by searching through a tree of possibilities[8]) takes time did not appear until the 1950s[9]. Furthermore, techniques to record the neural activity of behaving rats did not come into use until the 1970s[10], and neural ensembles were not accessible until the 1990s[11]. The decoding of neural activity at fast timescales was not possible until the 2000s[12–16]. These decoding operations enable the identification of neural correlates of imagination[17,18] even in non-human animals[19]. Perhaps the pausing behaviour identified as VTE really does reflect a search-and-evaluate process through imagined worlds.

In this Review, I present evidence that the neural processes that accompany VTE reflect a deliberation process in non-human animals. This article first defines deliberation algorithmically and compares it with procedural processes and other algorithms known to drive decision making in humans and other animals. I then review the behavioural and neurophysiological data and argue that these data support the description of VTE as reflecting deliberation. Finally, I address the multiple theories of VTE that have been proposed, concluding that the most likely explanation is that VTE is a behavioural manifestation of deliberative processes.

Defining deliberation

Deliberation is a process wherein an agent searches through potential possibilities based on a hypothesized model of how the world works and evaluates those hypothesized outcomes as a means of making decisions[20–24]. Deliberation depends on three steps: determining what those possibilities are, evaluating the outcomes and then selecting which action to take. Importantly, in humans, the deliberative search-and-evaluate process is thought to be serial, with individual options constructed as imagined concrete potential futures (a process known as 'episodic future thinking' or 'mental time travel')[20,21,23,25,26]. How options are compared against each other remains complex and only partially understood[27–29].

Search-and-evaluate processes depend on having a schema of how the world works that can be used to determine the consequences of actions[24,30–32]. In Tolman's terms, this schema is the so-called cognitive map, which was originally more cognitive than map[3,10,33,34]. Although many rodent studies treat the cognitive map as spatial, most theories assume that similar processes can occur through non-spatial schemas in all mammals[35–37]. Because VTE is specifically a rodent behaviour that manifests in a spatial context, only spatial analyses are reviewed here. Nevertheless, the neural processes underlying VTE are likely to reflect a more generally applicable deliberative process beyond those seen in the spatial examples used here.

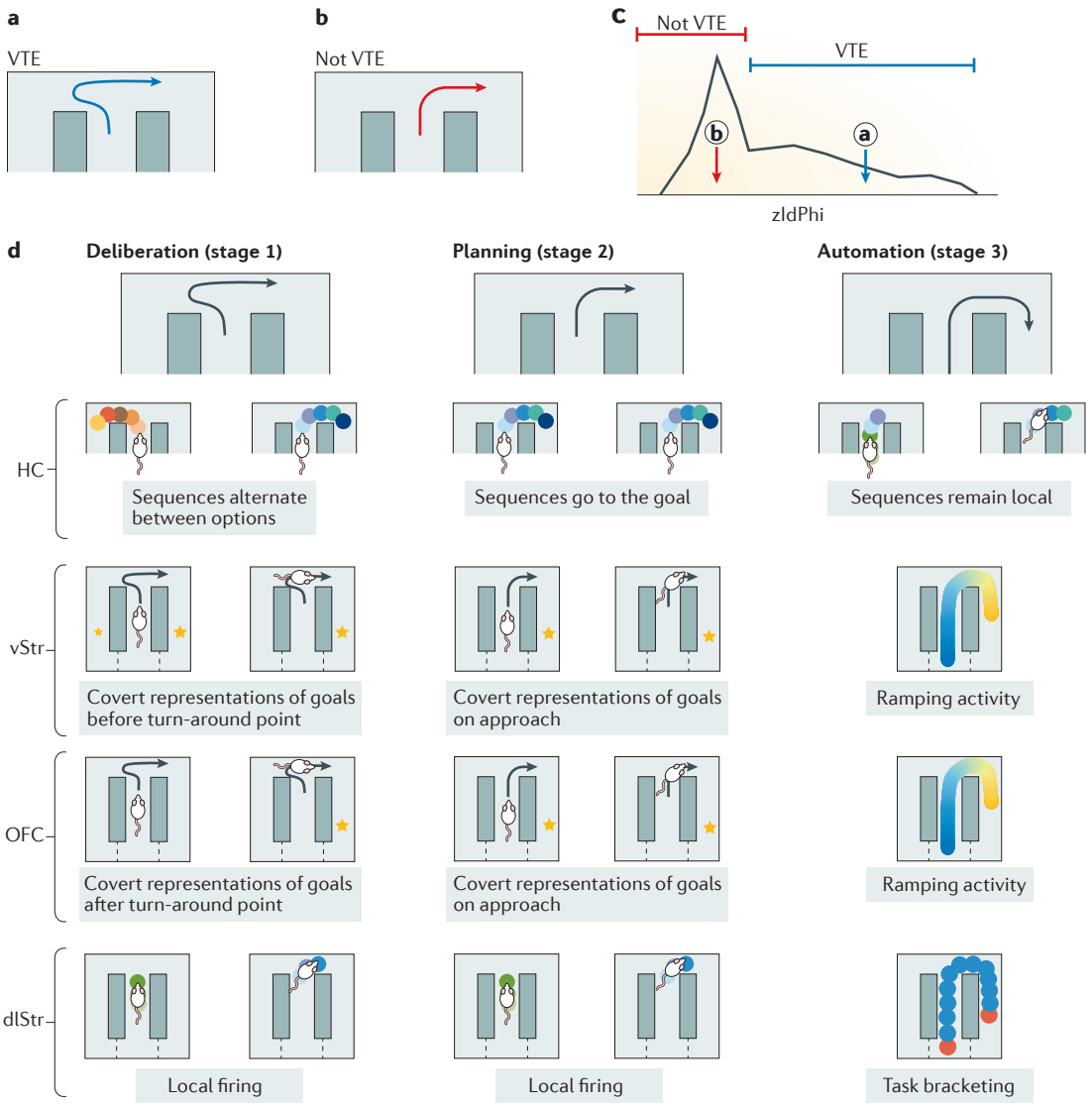


Figure 1 | **Vicarious trial and error**. The rodent behaviour originally termed 'vicarious trial and error' (VTE) by Muenzinger and Gentry[1] and proposed as a prospective imagination of the future by Tolman[2,3] is fundamentally a behavioural observation of pausing and reorienting. **a** | Example of pause-and-reorient VTE behaviour from T-choice experiments. **b** | Example of non-VTE behaviour: the animal orients towards only one trajectory at the choice point and continues along this trajectory. **c** | It is possible to quantitatively differentiate VTE laps from non-VTE laps using the zIdPhi measure. zIdPhi measures the integrated absolute angular velocity of the orientation of motion across the choice point; thus, it is high when the animals show reorientation behaviours and low when the animals simply pass through the choice point without a reorientation behaviour. zIdPhi shows a central peak (non-VTE) and a long right-skewed tail (VTE)[52,66,68,89]. On the panel, the circled 'a' shows where the path shown in panel **a** would fall on the zIdPhi score and the circled 'b' shows where the path shown in panel **b** would fall. **d** | VTE has been proposed to reflect prospective imagination and evaluation of the future. Accordingly, the shift from a flexible, deliberative decision-making process that requires this imagination to a habitual, procedural decision-making process[53,54,56] can be divided into three stages (top row): deliberation (in which VTE is expressed), planning (during which VTE is diminished) and automation (when the animal expresses no VTE and instead releases an action chain). Neurophysiological data suggest information processing consistent with this hypothesis. In the hippocampus, sequences of firing of hippocampal place cells represent sweeps ahead of the animal, serially exploring the paths towards the potential goals. As behaviour automates, these sweeps transition from going in both directions to going only in one direction, and then vanish[14]. Although specific neural correlates of prefrontal cortex (PFC) firing remain unknown, PFC–hippocampus interactions are increased during VTE[101,112,157] (not shown in the figure). Reward-related cells in the ventral striatum (vStr) transiently fire before the turn-around point in VTE. This early firing disappears as behaviours automate[129,132]. Reward-related cells in the orbitofrontal cortex (OFC) transiently fire after the animal commits to its decision. This firing appears earlier as behaviours automate[66,129]. Cells in the dorsolateral striatum (dlStr) do not show extra activity during VTE[54], but develop task bracketing (that is, they show increased activity at the start and end of the maze) as behaviour automates and VTE disappears[53,56].

Deliberative decision making in humans. Much of the extensive literature on human decision making differentiates between processes of deliberation and those of judgement[20,38,39], a distinction that parallels the difference between choosing-between-options and willingness-to-pay experimental paradigms in non-human animals[29]. Although these differences are beyond the scope of this Review, mounting evidence in the human literature indicates that processes underlying deliberative decision making depend on episodically imagined futures[21,25–27,40], such that more-concrete options are easier to imagine and draw decisions towards them[27,41,42]. In humans, the search-and-evaluate process is hypothesized to involve interaction between the prefrontal cortex (PFC) and temporal lobe structures to create imagined episodic futures[26,41,43–46]. In humans and non-human primates, medial striatal structures (caudate) may be involved in deliberative processes, whereas the more-cached, automated processes depend on lateral striatal structures such as the putamen[47,48].

Procedural versus deliberative decision making. Current theories of decision making suggest that decisions arise from a complex interaction of multiple action-selection processes[22,24]. Although a full description of this interaction is beyond the scope of this article, studies of VTE have examined behaviour during processes underlying both deliberative and procedural decision making. In procedural action selection, animals automate their behaviour when repeated actions reliably achieve goals, by developing cached action chains that can be released at appropriate times[4] but that, once initiated, tend to run to conclusion[49]. Because these action chains are inflexible once learned, they tend to be learned slowly, particularly in comparison to the cognitive map that forms the basis for the mental search process (knowing the cognitive map does not force the selection of a specific action)[10,24,34,50].

On the Tolman–Hull plus maze and similar T-choice mazes (BOX 1), VTE occurs when animals express deliberative strategies and vanishes as animals automate their response using procedural strategies[51,52]. However, VTE is also expressed on these tasks as procedural strategies are developing[14,51–53], which may suggest the presence of different underlying neural processes during the development of procedural strategies (before automation)[53–56].

VTE as a model of deliberation. The hypothesis put forward in this Review is that VTE reflects the indecision underlying deliberation. Predictions can be derived from this hypothesis about the timing of VTE, the relation of VTE to learning and reward-delivery contingencies, and the neurophysiological processes that should co-occur with VTE behaviours. See BOX 2 for a list of the predictions that follow from this hypothesis and that are supported by the evidence discussed in the following two sections.

Deliberative decision making
A process in which one imagines potential future outcomes (serially and individually) and then selects an action to get to that specific future outcome.

Procedural decision making
A process in which one learns an action chain and the ability to recognize the situations in which to release it. Performance is rapid, but is usually learned slowly, and is inflexible once learned.

Behavioural predictions

The key hypothesis of deliberation as a search-and-evaluate process is that animals are mentally searching through a schema of how the world works. In order to search through schema, a sufficiently complete model of the world is required that can predict the consequences of one's actions [24,57,58]. The hypothesis that VTE reflects deliberative but not procedural processes suggests that VTE should occur when rats depend on flexible decision-making strategies, which tend to occur early in learning[10,34] and after changes in the contingency of reward delivery[59,60]. Moreover, the hypothesis that VTE reflects the indecision underlying deliberation suggests that it should occur in particularly difficult choices and should be reduced when choices are easy[61].

VTE from deliberation to automation. Behavioural testing on the Tolman–Hull plus maze starts with habituation of the animal to the maze itself for several days, after which animals are trained to turn left from the south arm to the west arm. VTE occurs during the early learning phase on this task[51], while animals are showing deliberative behaviours[62,63]. Thus, VTE occurs when rats know the shape of the maze but not what to do in the maze. Animals that maintain the deliberative strategy late in learning continue to show VTE, whereas animals that switch to procedural strategies do not[51]. When variants of this task are constructed such that only the deliberative strategy will lead to a reward, animals continue to show VTE throughout the task[51,52]. On a plus-maze variant in which rats either had to go to a place from multiple starting points (requiring a flexible action) or had to take a specific action (for example, turn left) from any starting point (thus allowing automation), VTE occurred when the reward was delivered based on which place the animal went to (that is, depending on flexible, deliberative strategies), but vanished when animals automated the component where reward was delivered based only on the action taken (that is, following a procedural strategy)[52].

VTE and changes in reward contingency. On T-choice tasks[14,53,54,56,64–66] (BOX 1), VTE occurs at high-cost choice points. On these tasks, rats know the general parameters of the task (for example, run through a central path, turn left or right for food and return to the loop) but not the specific reward-delivery contingency (which could change daily). For example, rats that were trained to turn left or right at the end of the central track for food but that did not know which way was going to be rewarded on a given day expressed VTE on the early laps on each day[14,54]. VTE also increases when animals encounter a change in the relationships between actions and reward[56,65,66]. VTE disappears as the behaviour automates and then reappears with the change in reward-delivery contingency. Similarly, VTE reappeared on the cued-T task when the reward was devalued[53], and on the plus maze when the task switched between needing a place- and a response-based strategy for reward[52].

VTE and the difficulty of choice. On the multiple-T task, VTE behaviours are only expressed at the final (difficult) choice point, but not at a topologically similar control point on one of the simpler Ts[54], suggesting that simple choices do not produce VTE. On the restaurant row task (BOX 1), animals make sequential foraging decisions about whether to wait for a reward offered at a given

delay. VTE behaviours do not appear when the rat is faced with very short delays (which are definitely worth staying for) or really long delays (which are definitely worth skipping), but are preferentially expressed when the economic offer is at the animal's decision threshold[67], which is presumably the difficult choice.

Importantly, VTE does not simply occur only in situations in which there are two equal choices between which a decision must be made. One study trained rats on a spatial delay-discounting task[68] (BOX 1), in which rats decide between two differently delayed and differently rewarding options. The rats can change the delays to the options by preferentially selecting one side or the other and, once the two sides are equally valued, the animals alternate between the two sides, which holds the delays constant. In this task, VTE occurs as the delays are changing and vanishes when the options become equally valued and the rats respond with a stereotyped alternating behaviour.

VTE and behaviour: conclusions. These data suggest that VTE occurs when animals are faced with difficult, changing choices on tasks in which they understand the shape of the world. However, if it is possible to automate the behaviour (when the reward-delivery contingencies are stable), then VTE disappears as behaviour automates.

The neurophysiology of VTE

On the basis of the hypothesis that VTE reflects deliberation and disappears behaviourally as animals begin to use procedural decision-making systems, several predictions about the neurophysiology accompanying VTE can be made (BOX 2).

Deliberation entails a search-and-evaluate process, in which one imagines the consequences of an action and then evaluates that outcome in light of one's goals. This is a form of mental time travel, or episodic future thinking: placing oneself into a future situation[25,69].

Box 1 | **Tasks**

Vicarious trial and error has been explored in rodents using several spatial tasks (see the figure).

Tolman–Hull plus maze
Rats are trained to turn left from the south arm to the west arm of a plus maze. The animals then start a trial from the north arm, allowing them totake a different action (to reach the same location) or the same action (leading to a different location)[51,62,88].

Place or response plus maze
Rats have to proceed to a place or follow a procedural action to obtain a reward[52].

Multiple-T task
Rats are trained to run through a central track and to turn left or right for food. On each day, the central track has a different shape and the side on which the final reward is placed can change, but both variables generally remain constant within a day[14,54].

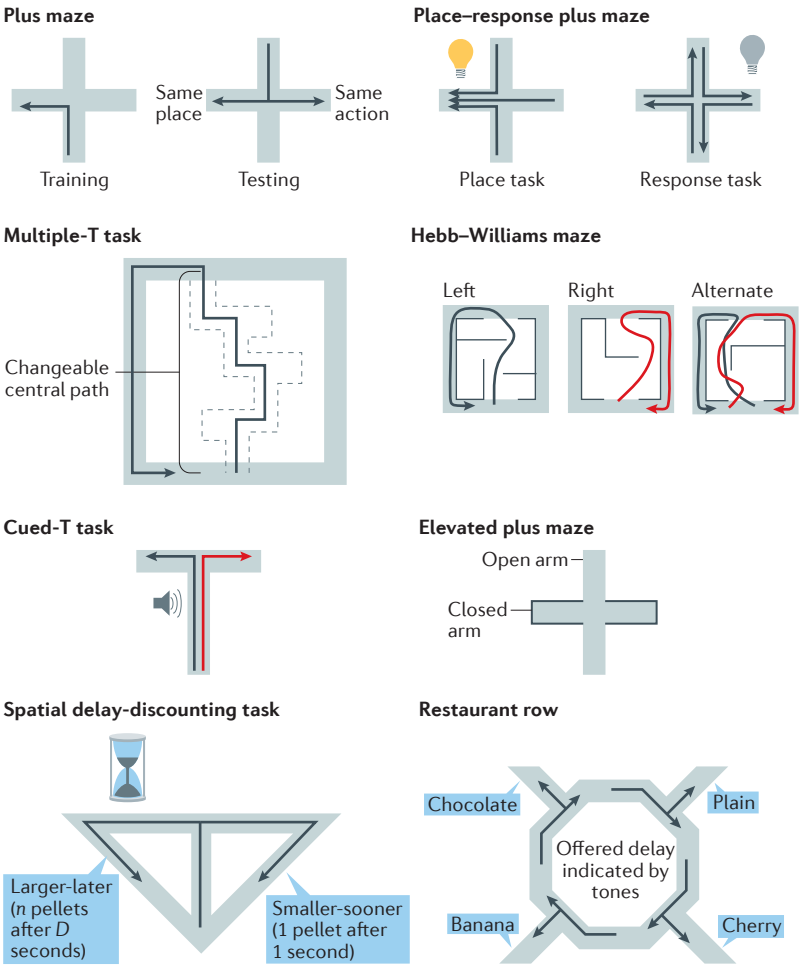
Left, right or alternation choice tasks
Variants of the multiple-T task in which the track or reward-delivery contingency changes part of the way through a day's experimental session[65,66,72]. In the Hebb–Williams maze version, the central path is constructed of walls rather than additional T choices[56].

Cued-T task
Rats or mice are trained to turn left or right at a T intersection where the correct choice depends on a sensory cue provided on the stem of the T[14,53,55,140].

Elevated plus maze
Rats or mice explore a plus maze with two closed and two open arms. The open arms are less protected and more-anxious animals tend to avoid them[153] (BOX 4).

Spatial delay-discounting task
Rats are trained to turn to one side of a T for a large reward that is delivered after a delay or to the other side for a smaller reward that is delivered quickly. The delay is adjusted based on the rat's decisions, allowing the rat to balance the delay against increased reward on the larger–later side[68].

Restaurant row task
Rats or mice are trained to run past a repeating sequence of reward options. As the animal runs past an option, a sensory cue informs it of the delay it would have to wait before receiving the reward. At each offer, the animal is given the option to wait out the delay or to skip the option and try the next option, allowing the assessment of how much the animal values that option (that is, through measuring how long it is willing to wait for a given option)[67].



Box 2 | **Predictions**

The hypothesis that vicarious trial and error (VTE) behavioural events reflect an underlying deliberation process generates several testable predictions, which the data reviewed in this paper suggest are true.

Behavioural predictions

- VTE should occur in situations in which the rat knows the shape of the world, because deliberation depends on being able to predict the consequences of one's actions.
- VTE should decrease as animals automate their behaviours when reward-delivery contingencies are particularly stable.
- VTE should increase when reward-delivery contingencies are variable or change.

Neurophysiological predictions

- During VTE events, there should be neural representations of future outcomes and their evaluations.
- Those neural representations should encode the multiple outcomes serially.
- Neural representations of future outcomes should involve the hippocampus and prefrontal cortex (PFC)–hippocampus interactions.
- Neural representations of valuation should involve the ventral striatum and the orbitofrontal cortex.
- Neural representations should shift from PFC–hippocampus interactions to dorsal striatal processes as behaviours automate.

Studying mental time travel in non-human animals requires a methodology that can detect representations of that imagined future[19], and just such a methodology has been developed that is based on the observation that imagining something activates similar representational systems to perceiving that thing[17,18] (BOX 3). This methodology can be used to explore how different neural circuits contribute to the search-and-evaluate process that underlies deliberation, and thus shows that mental time travel itself uses the hippocampus, in a process that is probably initiated by the PFC, and that the outcomes of each possible option are evaluated in the ventral striatum (also known as the nucleus accumbens) (see below).

By contrast, procedural decision making entails the performance of an appropriate action chain in a single situation. Thus, structures involved in procedural decision making should slowly develop representations that relate specific actions to specific reward-delivery contingencies, and should directly reflect the 'release' of an action chain. The dorsolateral striatum has been shown to develop such representations as animals automate their behaviour (see below).

Hippocampal sequences. Normally, as a rat runs through a maze, each hippocampal place cell fires action potentials when the animal is in a specific location — in the place field of the cell[10,34]. During movement (for example, when running through a maze) and attentive pausing, place cell firing is organized within an oscillatory process that is reflected in the hippocampal local field potential (LFP) theta rhythm (6–10 Hz)[10]. Typically, theta is seen during movement, and other LFP patterns are seen during pauses; however, theta also occurs if the rat is attentive to the world during that pause, such as during freezing in fear after conditioning[70] or during VTE[10]. In each 140 ms theta cycle, hippocampal place cells fire in a sequence along the path of the rat, from somewhere slightly behind the rat to slightly ahead of it[15,71,72]. By examining these hippocampal theta sequences during VTE events at decision points on a T-maze task, it was found that these sequences proceeded far ahead of the rat, alternating serially between goal-related options[14].

Evidence from humans suggests that the hippocampus is both necessary for and active during episodic future thinking[25,26,45,73,74]. Together with the evidence described above that neural representations in the hippocampus project sequences forward along the path of the rat, this evidence suggests that the pause-and-look behaviour observed as VTE may indeed be a manifestation of deliberation, particularly on spatial tasks. Moreover, for the rat to be internally 'trying out' possible outcomes, the hippocampal place cell sequences should occur serially, alternately exploring the paths to each goal ahead of the animal (as opposed to representing a blend of all possible paths). This is exactly what occurs: when animals pause at the choice point, hippocampal sequences within any given theta cycle proceed along a single path towards a single goal, one at a time, first towards one goal and then towards the other[14,72,75]. Current experiments have only examined neural ensembles ranging over a small portion of the hippocampus (within a few millimetres). Recent work has suggested that the hippocampus and its associated structures may contain several potentially independent modules along the long septo-temporal axis[76,77]. Whether the hippocampal neural ensembles encode a unified representation of a 'sweep' in both the dorsal and the ventral hippocampus remains unknown. One can quantify how unified a representation is by measuring whether the activity of the neurons is consistent with the location decoded from the population, termed the 'self-consistency' of a representation[19] (BOX 3). During VTE, dorsal hippocampal neural ensembles encode self-consistent representations of each individual option, serially and transiently.

These hippocampal sweeps during VTE were locked to the hippocampal theta oscillation[14], suggesting that they may be special cases of theta sequences. During each theta cycle during normal navigation, hippocampal place cells fire in sequence along the current path of the animal[15,78,79]. Hippocampal theta sequences have been found to reflect 'cognitive chunks' (that is, continuous sequences of navigational paths) about goals and subgoals[72]. Moreover, these sequences have been found to represent the path to the actual goal of the animal, sweeping past earlier options at which the animal was not going to stop[75].

Hippocampal LFPs: theta and LIA. In contrast to the theta rhythm, the more broad-spectral LFP known as large-amplitude irregular activity (LIA) is punctuated by short 200 ms oscillatory events called sharp-wave ripple complexes (SWRs) that contain power in frequencies of approximately 200 Hz[10,80]. SWRs that occur just before an animal leaves a location where it has been resting often include hippocampal SWR sequences running forward towards the goal[16,81], and thus it has been hypothesized that there may be SWRs at the choice point during VTE. However, the rodent hippocampal LFP shows theta, not

Local field potential
(LFP). Low-frequency voltage signals reflecting neural processing. In the hippocampus, the LFP is marked by two contrasting states: theta and large-amplitude irregular activity.

Hippocampal theta sequences
Sequences of firing of hippocampal place cells within a single theta cycle, generally proceeding from the location of the rat forward towards potential goals. Also called a 'hippocampal sweep'.

Hippocampal SWR sequences
A sequence of firing of hippocampal cells within a sharp wave ripple complex (SWR). Originally referred to as 'replay' (because early observations identified sequences repeated in order), but now known to include other sequences including backwards along the path of the animal or along unexplored shortcuts and novel paths.

LIA, during VTE[14,72,82], and there is little to no power in the SWR band during VTE[14]. Although there are intra-hippocampal oscillatory events that occur shortly before the turn-around point in VTE, these events show power at frequencies in the high-gamma range (~100 Hz), well below the SWR range[82]. The relationship between SWRs that occur before the journey starts[16] and theta sequences that alternately explore goals during VTE[14] remains unknown.

Three stages of hippocampal sequences. Johnson and Redish[14] first saw these VTE-related sequences on the simple multiple-T task, in which rats had already learned the general structure of the task (that is, run through the central track and turn left or right for food), but not the correct decision for that day (because reward-delivery contingencies changed from day to day). Rats that were familiar with the general structure of the task learned the correct choices quickly[14,54,83], which suggests that they were finding the daily parameters to fit a learned schema[84]. Because rats ran the same path repeatedly through the day, they could eventually automate their behaviour, leading to highly stereotyped paths with experience[54].

By decoding the hippocampal representations while the rat paused at a choice point (while the rat expressed VTE[54]), it was found that hippocampal representations serially explored both alternatives on early laps, sweeping first to one side and then the other[14]. With experience, VTE diminished but representations continued to sweep ahead of the animal; however, in these later laps, sweeps only went in one direction (the direction eventually chosen by the animal). With extended experience, VTE was no longer expressed, the path of the animal became very stereotyped, and hippocampal activity represented the local position of the animal, with little forward components at all.

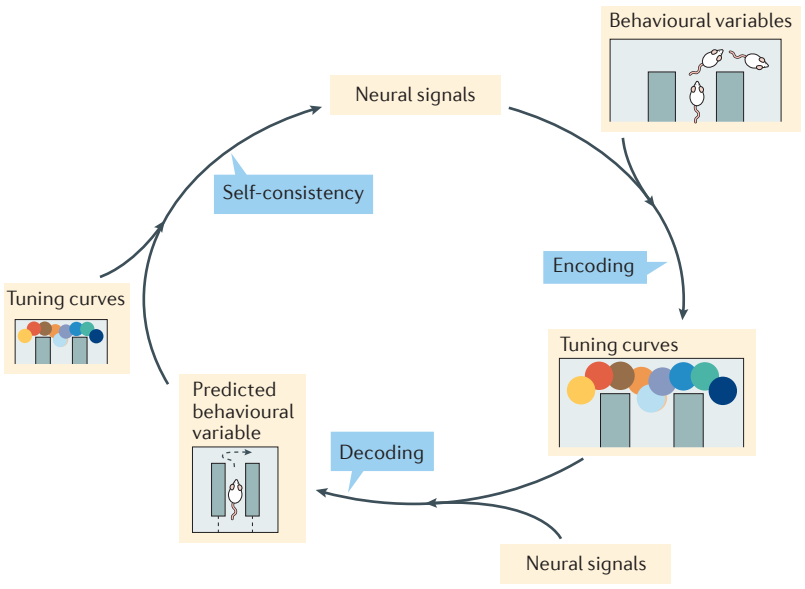
I suggest that these three components of hippocampal representation entail three stages of behavioural automation. The first stage is one of indecision, in which rats know the structure of the world but need to vicariously imagine the alternatives to determine what they want to do. The second stage is one of planning, in which rats know both the structure of the world and what they want to do but, using a deliberative process, continue to run the search through only one option (checking to ensure that option is the one they want). The third stage would then be automation, in which rats no longer search the future, but rather simply execute an action chain, with the hippocampus only tracking the immediate present (FIG. 1).

Hippocampal manipulations and VTE. Hippocampal lesions decrease the expression of VTE, particularly during learning, investigatory and exploratory stages, but these same lesions can increase the expression of VTE in later stages[85–87]. Ageing, which disrupts cognitive spatial strategies through effects on the hippocampus[88], has a similar effect of increasing VTE during later stages

Box 3 | **Measuring imagination**

In humans, imagination of a thing activates the same representational systems that are active during perception of that thing[17,18], and decisions that depend on an explicit knowledge of how future situations transition to other future situations include prospective representations of those future cues[98]. This suggests that mental time travel in non-human animals can be detected by decoding representations from patterns of neural activity and looking for times at which the neural activity decodes to information not immediately present — that is, when it decodes to other places or other times.

To measure representations of other places and other times, we follow a representational loop of encoding, decoding and re-encoding[19] (see the figure). The encoding step ascertains the relationship between neural firing and a behavioural variable (such as the position on the maze). The decoding step reveals the information represented by a specific firing pattern (such as a firing pattern at a particular moment in time). The third, re-encoding step measures how well the decoded information predicts the observed spike train pattern. Self-consistency is a quantitative measure of the encoding–decoding–re-encoding loop, taken as how consistent the activity patterns of cells is with the information hypothesized to be represented at a moment in time[154]. Hidden within the encoding step is an assumption about the variables that the cells are encoding (for example, that hippocampal cells encode spatial information). Incomplete assumptions about the variables that the cells are encoding will decrease the self-consistency of the representation[19,154]. Sometimes, one needs to add a 'covert' variable (not shown on the figure), such as the imagined location of the animal, in order to correctly explain neural firing. The cells fire action potentials outside their expected tuning curves (that is, in conditions besides those to which the cells normally respond); if those spikes contain structure (for example, sequences, consistency with other simultaneously recorded cells or information about future plans), then this extra structure can be interpreted as representing cognitive information about other places or other times[19].



Task bracketing
A phenomenon observed in dorsolateral striatal neural ensembles, in which the cells show increased activity at the beginning of an action chain sequence.

of tasks[89]. In addition, in the elevated plus maze (a task used for measuring levels of anxiety; see BOX 4), hippocampal lesions disrupt the pausing behaviour that rats exhibit before venturing on to the open arms. It may be that hippocampal lesions reduce the expression of VTE, but also prevent animals from shifting to the planning and automation stages as easily. As noted above, on tasks in which rats show a shift from a flexible behaviour to an automated behaviour, such as the multiple-T task[54], the Tolman–Hull plus maze[51,62,88] or the spatial delay-discounting task[68], VTE typically appears during early performance along with flexible behaviour and vanishes as the behaviour automates. On these tasks, hippocampal lesions decrease the VTE that occurs early (suggesting a problem with deliberative systems), but then these hippocampal lesions also disrupt this change in VTE with experience, leading to an increase in VTE, particularly on later laps (which should be more automated). Whether this disruption in the normal decrease in VTE also disrupts processes associated with automation (such as the development of task bracketing in the dorsolateral striatum; see below) remains unknown.

Intra-hippocampal pharmacological manipulations can also change the expression of VTE. The cannabinoid agonist CP55940 has three notable effects on behaviour and hippocampal representations, whether given systemically or intra-hippocampally. First, it disrupts the ability to hold information across temporal gaps between trials[90]. Second, it disrupts the phase-related firing of place cells[91]. Third, it markedly increases VTE[90,92]. Given that phase-related firing of place cells is a reflection of the underlying place-cell theta sequences[15,78], the cannabinoid-induced disruption of phase-related firing is likely to be reflective of the disruption of the theta sequences and thus indicates a disruption of the forward sweeps. A plausible interpretation of these disruptions is that CP55940 hinders the ability of the hippocampal sweeps to produce a self-consistent representation of the episodically imagined future, leaving an animal in the indecision stage, with continued expression of VTE.

In contrast to CP55940, reduction of the tonic release of noradrenaline by clonidine decreases VTE[93]. It is not known what effect clonidine has on theta sequences, but new data suggest that it limits the search process to a single path, even during VTE events[94]. This mentally searched path is the one that gets chosen by the rat[94]. A plausible explanation for the reduction in VTE following clonidine exposure is that clonidine drives the animal into the planning stage, in which it mentally explores only a single option and thus does not spend as much time expressing VTE.

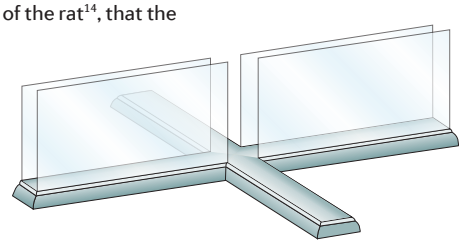
Beyond the hippocampus. The neurophysiology of hippocampal representations during VTE[14] suggests that the hippocampus enables the representation of the imagined future outcome required by the search component during the VTE process. Given the known interactions between the hippocampus and cortical representation systems[34,84,95], it is likely that a full representation of an episodic future would depend on both hippocampal representations of the cognitive map and cortical representations of the features in that episodic future[96,97]. This is consistent with observations that episodic future thinking in humans activates hippocampal and medial temporal lobe systems[26,73], and prospective representations of specific options activate cortical areas that can be used to decode those future options[98].

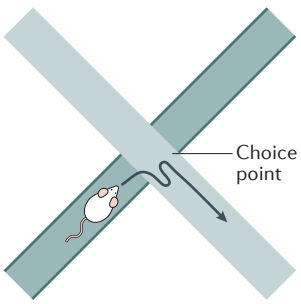
Because disrupting hippocampal representations increases VTE[87,90], it is unlikely that the hippocampus initiates the VTE process. I suggest, instead, that some other structure polls the hippocampus for that episodic future. Given the known influence of the rodent

Box 4 | **Anxiety**

A hallmark of anxiety is the inability to act owing to pausing to consider possibilities, and particularly the dangers inherent in those possibilities. A hallmark of rodent models of anxiety is an unwillingness to venture out into open spaces. For example, on the elevated plus maze (see the figure), anxious rats prefer to stay within the enclosed arms and do not like to venture out into the open arms[153]. This model seems to have construct validity in terms of involvement of similar brain structures as in human anxiety[155,156] and similar effects of anti-anxiolytics[155].

The central intersection point between the closed and open arms of the elevated plus maze forms a choice point (see the figure), and rats often show micro-choices that look very much like VTE events at this choice point[153,157]. There is increased interaction between the hippocampus and the prefrontal cortex during these events[157,158]. This pausing at this central intersection point has been referred to as a 'stretch–attend posture', during which the rat leans out and then returns back[159]. The stretch–attend posture occurs at times of conflict between approach to a reward (such as getting a food pellet) and avoidance of a threat (such as a dangerous predator-like robot)[160,161]. The stretch–attend posture has been hypothesized to be critical for risk assessment[153,159] and probably represents a VTE event between staying and going. Predictions of this hypothesis are that there will be hippocampal theta sequences during these stretch–attend postures that will represent trajectories ahead of the rat[14], that the hippocampus and the amygdala will show local field potential coupling[70], and that there will be evaluation-related representations in structures such as the ventral striatum and the amygdala[132,161].





prelimbic cortex on outcome-dependent decisions[99,100] and effects of its manipulations on goal-related activity in the hippocampus[101,102], the prelimbic cortex is a good candidate for the structure that is likely to be initiating VTE.

The PFC. Theories of the neurophysiology underlying human deliberative processes, particularly the initiation of imagination, suggest a role for the PFC in deliberation[26,45,73]. In the rat, the prelimbic and infralimbic cortices (together termed the medial PFC (mPFC)) are usually identified as homologous to the primate dorsolateral PFC[103,104], although the homology remains controversial[105].

Cells in the mPFC tend to represent task-related parameters and, on structured maze tasks, their activity tends to chunk the environment into task-related sections[106,107], analogous to the trajectory-related sequences seen in the hippocampus[72,75]. It is not known whether these mPFC neural patterns are directly related to the hippocampal sequences, but, interestingly, mPFC cell-firing patterns lock to hippocampal rhythms during episodic memory and decision-making tasks[108,109], and interfering with the prelimbic–hippocampal circuit (through optogenetic silencing or physical lesions of the nucleus reuniens) disrupts the processing of goal-related information in the hippocampus[102]. Disruption of the mPFC of rats impaired the ability of the animals to flexibly switch to novel strategies[110], but not after those strategies had been well learned[111], suggesting that the mPFC is necessary during those task aspects in which VTE tends to appear[52].

There is a strong, transient increase in LFP interactions between the hippocampus and the mPFC at the choice point during times when rats typically express VTE[101,112,113]. Several researchers have suggested that coherent oscillations (measured by paired recordings of LFPs in two separate structures) can increase the transmission of information between structures, allowing for transient changes in functional connectivity as a function of different task demands[114]. The functional relationships between the hippocampus and the mPFC are critical for the goal-dependent firing of hippocampal cells[102], and are disrupted by cannabinoids[115].

These studies suggest a model in which VTE reflects a deliberation process entailing an interaction between the mPFC and the hippocampus. I propose that the mPFC initiates the request, and the hippocampus responds by calculating the consequences of a potential sequence of actions (that is, identifying a potential future outcome). This interaction may explain why the hippocampal sequences are serial considerations of potential options — each request by the mPFC is searched through by a single theta sequence. If the hippocampus responds with poorly structured information (such as when under the influence of cannabinoids), the mPFC sends additional requests, leading to an increase in VTE. If the hippocampus responds assured of the consequences of the action (such as when under the influence of clonidine), the mPFC is satisfied and VTE is diminished. However, how the quality of the hippocampal representation is evaluated by the mPFC or other downstream structures remains unknown.

The ventral striatum and the OFC. It is not enough to know what will happen when exploring a potential path during deliberation — one also has to evaluate the consequences of that path, preferably in light of one's goals[57]. Both the ventral striatum and the orbitofrontal cortex (OFC) have been identified as key to evaluation processes in rats[116–118], monkeys[119–121] and humans[122–125]. Neurons in these structures respond to reward value[119,126–129], and reward-related cells in both the ventral striatum and the OFC fire transiently just before an animal initiates a journey to a goal[127,130,131]. Cells in these structures that exhibit large firing responses to reward also fire action potentials during VTE[66,129,132]. This suggests the presence of an imagination of reward during VTE events, potentially providing an evaluation signal.

On T-maze tasks, the pause-and-reorient behaviour of VTE has one (or more) very clear, sharp transition that can be identified quantitatively[54,66,68,82,129], and can provide the potential to identify the timing of neural reward representations relative to decision-making processes. Covert reward signals in the ventral striatum preceded the moment of turning around, whereas the covert reward signal in the OFC appeared only as the animal oriented towards its goal, after turning around[54,66,129].

These data suggest that the ventral striatum provides evaluative calculations during the decision-making process itself (before the animal stops and reorients towards its goal), whereas the OFC provides expectation information only after an animal commits to its decision. This hypothesis is consistent with recent work examining the timing of choice-related value representations in primates, comparing when evaluation signals in neurons recorded from the ventral striatum and ventral prefrontal cortical areas (which are probably homologous to the rat OFC) represent the value of the selected choice[121]. It is also consistent with work finding that the rat OFC exhibits outcome-specific representations of expected rewards after decisions[67,133].

The dorsolateral striatum. As noted above, when repeated actions reliably achieve goals, animals can use procedural decision making, which depends on an algorithm of releasing action chains after recognizing the situation[49]. In a sense, the animal has learned to categorize the world (by developing a schema) and acts appropriately given those categories[24]. A key structure in the categorize-and-act process is the dorsolateral striatum. Lesions and other manipulations that compromise the function of the dorsolateral striatum drive behaviour away from procedural decision making and towards deliberative decision making[62,63]. Theoretically, the information-processing steps involved in procedural decision making are considerably different from those involved in the search-and-evaluate, deliberative system[22,50,54].

Structures involved in procedural decision making should slowly develop representations of situation–action pairs, but only for the cues that provide reliable information as to which actions should be taken to achieve a reward. In contrast to the ubiquitous spatial representations of hippocampal place cells[10,34,134,135], structures involved in procedural learning should only develop

Covert reward signals
Signals reflecting imagined representation of reward, as detected from the patterns of activity of reward-associated neuronal ensembles during non-rewarded events.

spatial representations when knowing where one is in space informs which actions one should take. Neural ensembles in the dorsolateral striatum only develop representations of those action–task pairs that reflect cues and task parameters that are informative about the action[55,136,137]. For example, dorsolateral striatal representations develop spatial correlates on the multiple-T task, but not on an analogous non-spatial task even when that non-spatial task occurs on a large track-based maze[83,137].

Moreover, dorsolateral striatal ensembles should not show activity that is representative of transient mental sweeps into future outcomes (as the hippocampus does). Indeed, they do not; even during VTE events, and even on tasks in which dorsolateral striatum cells do show spatial representations[54,56]. Nor should reward-related cells in the dorsolateral striatum show transient activity carrying covert information about reward (as the ventral striatum and the OFC do) — again, they do not; although a subset of dorsolateral striatum cells do show reward-related activity[53,56,83], these reward cells do not show a transient upregulation of firing during VTE events[54]. Thus, neural activity in the dorsolateral striatum during decision-making tasks is more consistent with an involvement of this structure in procedural decision making than in deliberative decision making.

As behaviour becomes 'ballistic' and stereotyped, dorsolateral striatal neural ensembles transition their firing to the start and end of the journey, with decreased firing in the middle[53,56,138–141]. This phenomenon has been termed task bracketing, and it diminishes quickly if a change in the reward-delivery contingency forces animals back into using deliberative strategies[53,56]. Task bracketing is negatively correlated with the presence of VTE on a lap-by-lap basis[53,56], supporting the hypothesis that VTE is indicative of decisions being made by the deliberative system, whereas task bracketing is indicative of the decisions being made by the procedural system. Interestingly, task bracketing remained after reward devaluation[53] and after switching the signalling cue to a novel modality[142], but disappeared during extinction[140] and after a reward-delivery contingency reversal that required a change in action in response to the same set of cues[56]. These findings may suggest that the hypothesized relationship between reward evaluation or devaluation and deliberative search processes[35] may require a more complicated process model[42,143] that separates one-step and multistep processes.

Neurophysiological conclusions. The neurophysiology summarized above is consistent with the hypothesis that there are (at least) two decision-making systems based on different algorithms — a deliberative system based on search-and-evaluate processes and a procedural system based on the release of learned action chains in response to situation categorization. It is also consistent with the hypothesis that VTE is a behavioural reflection of the deliberative process. The search process arises from an interaction between the PFC and the hippocampus, with pre-decision evaluation in the ventral striatum, and post-decision evaluation in the OFC. In the rat, the infralimbic cortex may have contrasting roles to the prelimbic cortex, instead affecting learned, automated procedural strategies[53,99]. The data reviewed above suggest that the dorsolateral striatum is a critical component of the procedural system; it learns the associated action chains that should be taken in rewarded situations. There is evidence that the dorsomedial striatum may have roles in the deliberative component[55,56,63,144–146], and that there may be differences between the functions of the anterior and posterior portions of the dorsomedial striatum[55,56,63], but the specific roles of the dorsomedial striatum in decision making remain incompletely explored.

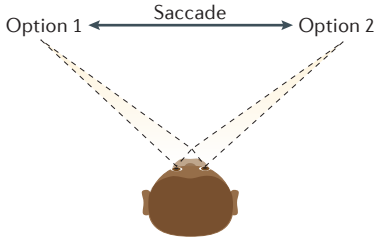
Theories of VTE

The term 'vicarious trial and error' as originally suggested by Muenzinger, Gentry and Tolman[1–3] is fraught with meaning, implying imagination, mental time travel and a cognitive search-and-evaluate process. In effect, they were proposing that rats were deliberating over choices. The neurophysiological data reviewed above suggest that this is actually an accurate description of the behaviour. Over the years, alternative accounts of VTE have been proposed, suggesting that perhaps VTE was simply a means of gathering sensory information or that it merely reflected indecision during learning. However, further examination of these two alternatives suggests that VTE really is an example of using working memory to explore mental possibilities — that is, to vicariously try out alternatives.

Is VTE a means of gathering sensory information? As a behavioural process, VTE entails orientation towards a path to a goal, or perhaps even orientation towards the goal itself. Because of this, early theories of VTE suggested that rats were merely focusing sensory attention on cues[1,147]. This simple hypothesis is, however, untenable, because the proportions of VTE change as behaviour changes, even when sensory cues are held constant[14,67,68,87]. Certainly, there are examples of pause-and-reorientation behaviour in both rats (looking like VTE) and primates (BOX 5) that have

Box 5 | Primates

This Review has concentrated on the rat because vicarious trial and error (VTE) behaviour has been both best defined and most extensively studied in this organism. However, in certain visuospatial tasks in which options are separated on a visual screen, humans and non-human primates do show a similar process termed 'saccade–fixate–saccade' (SFS), in which the subject looks from one option to another[61,162,163] (see the figure). Thus, just as VTE in the rat entails alternate orientation and reorientation towards goals, SFS may entail an alternate orientation and reorientation towards a target. This connects SFS to the extensive primate (both human and non-human) literature on visual search[36,164]; however, whereas visual search is usually used to refer to situations in which perception is difficult but the decision — once the correct stimulus is perceived — is easy, SFS is usually observed in situations in which the perception is easy, but the decision is hard, such as when given a choice between two similarly valued candy bars[61]. As with the rats and VTE, amnesic patients with severe hippocampal dysfunction showed disruptions in these SFS revisitation processes[162], which are believed to depend on prefrontal–hippocampal interactions[45].



Integration to threshold
A psychological theory of decision making whereby one accumulates evidence for one decision over another; when the evidence for one decision reaches a threshold, the decision is made.

information-gathering components. However, hippocampal place-cell sweeps do not always reflect the orientation of the rat[14], suggesting that the sweeps reflect an internal process of vicarious exploration rather than a means of gathering external sensory information. Although it is true that rats have a very wide field of view, VTE has been observed in walled-maze tasks in which the animal cannot see the goal[56,87], so any sensory-information-gathering theory would have to depend on only the first part of a journey. Moreover, VTE-like reorientations observed in sensory-information-gathering situations do not depend on hippocampal integrity, unlike VTE-like reorientations in goal-directed behaviours[86]. VTE is more likely to be a reflection of an internal process[50,58], perhaps as the halted initiation of an action.

Is VTE exploration of novel environments? On initial exploration, rats do show specific exploration-related behaviours — particularly pauses during which rats tend to show visual exploration of an environment by rearing up on their hind legs and swinging their head back and forth from side to side[10,148]. Although VTE often does include the head swinging from side to side, this behaviour reflects the orienting and reorienting towards paths and does not include rearing on the hind legs. Exploratory behaviours occur on novel mazes on which rats have very limited experience, whereas VTE occurs on tasks in which the animal has lots of experience but requires specific parameters for a given day. Moreover, rearing behaviours occur throughout the space of the maze, whereas VTE occurs specifically at difficult choice points. On a circular track, animals expressed pause-and-rearing behaviours randomly throughout the track[149], whereas on the multiple-T task, VTE behaviours were only expressed at the final (difficult) choice point[14,54], and on the restaurant row task VTE behaviours were preferentially expressed at entries onto the reward spokes, particularly when the economic offer was at the animal's decision threshold[67]. On the spatial delay-discounting task, more VTE occurred during the so-called titration phase (while rats adjusted the delay) than during the investigation or exploration phases[68,87]. Neurophysiologically, VTE is associated with hippocampal sweeps of representations of paths towards a goal and do not produce long-term changes in the place fields of the cell[14]; by contrast, exploratory pause-and-rear behaviours recruit hippocampal cells to develop new place fields[149]. Together, this evidence suggests that VTE is more likely to reflect a form of exploration within an internal, mental space of possibilities than exploration of the external environment.

A search through mental information space. One of the discoveries of the computer revolution is that information can be hidden within complex data structures such

Box 6 | **Open questions**

Although we have made a lot of progress in identifying the neural mechanisms underlying vicarious trial and error (VTE) behaviours, confirming many of the original implications of the term 'vicarious trial and error' (an explicit search process representing future options), several key open questions remain.

How does the process decide which potential futures to search?
Mental construction of potential future outcomes is a form of a search process[23,36], and thus requires a recall process that must be modulated by one's goals. Cueing expected outcomes can guide preferences towards that outcome, but it is not yet known whether those preferences also drive mental search representation to focus on paths to that outcome[35,41,44,165–168].

The information represented within hippocampal sequences is correlated with the orientation of the rat during VTE, but does not track it on a one-to-one basis[14,72]. Similarly, hippocampal sequences during goal-related navigation include increased representations of the chosen goal but are not uniformly targeted towards the chosen goal[16,75]. This suggests that VTE probably reflects the process of searching but is not necessarily oriented towards the goal itself. How these two processes (the neurophysiology of search and the orientation of the rat) are related remains incompletely explored.

How are the options evaluated?
In humans, evaluation of episodically imagined futures during deliberation seems to be based on applying perceptual valuation processes to the imagined outcomes[21,27,28,169]. Perceptual and deliberative valuation processes in humans and other animals activate and depend on similar neural systems[28,122,124,165]. Evaluative steps during VTE in rodents activate the ventral striatum and the orbitofrontal cortex[66,129,132]. In humans, the evaluation step depends on current emotional processes[20,27,28]. In rats, devaluation of a reward reduces the performance of actions that lead to that reward during early learning and when flexible behaviours are required (because reward-delivery contingencies are variable); however, after automation, reward devaluation has no such effect on the performance of these actions[99]. This discrepancy suggests that during deliberative processes there may be an evaluation step that is based on current needs and that is not used during procedural processes[57,165]. Whether this step works through a homologous process to human evaluative processes remains controversial[143,166,170].

How is the action selected?
Deliberation is a decision-making process[24]. Studies have so far failed to find any relationship between the directions being represented during VTE and the eventual choice of the animal in normal animals[14,129]. The saccade–fixate–saccade processes in primates (BOX 5) have been argued to reflect integration of value of a given choice[61], which accesses the psychological decision theory of integration to threshold[171]. This model has extensive psychophysical and neurophysiological support in perceptual decisions in which one needs to determine how to perceive a complex stimulus[172,70]. For example, when a subject is faced with a set of dots on a screen that are moving mostly randomly and asked to determine the average non-random direction hidden within the random population, representations of the alternative possibilities develop slowly, integrating until they cross a threshold[172]. However, neural ensemble recordings in the ventral striatum and the orbitofrontal cortex of rats have failed to find integration-to-threshold signals during VTE in either structure[129]. Behavioural data on some tasks indicate that VTE is increased when decisions are difficult[67], which is consistent with the integration-to-threshold hypothesis (more difficult decisions require more integration time), but in other tasks VTE is increased only when the animal is using a deliberative strategy. On the spatial delay-discounting task, VTE occurs when values are very different from each other and disappears when the values become equal[68] (presumably because the rat can use a procedural strategy at that point on this task). Although the action-selection process that must end deliberation remains unknown at this time, the evidence strongly supports the idea that VTE is reflecting the search-and-evaluate process that underlies deliberation.

Visuospatial scratchpad
A component hypothesized to underlie working memory processes, in which neural circuits normally used for perception are used to hold imagined information for processing.

that accessing that information can require computation and time[8]. Similarly, the information about which rewards are available, and where and how to get to them may be stored in an animal's brain, but determining what the consequences of one's actions are and whether those actions are the right choice at this time can require a search process through one's knowledge about the world.

Interestingly, early theories of working memory suggested that working memory allowed information to be explored internally to determine previously unrecognized aspects hidden within one's knowledge[150]. This exploration was hypothesized to arise from prefrontal areas[45,151], using sensory and other representational systems as a visuospatial scratchpad[18,150]. The deliberative explanation for VTE derived in this article also suggests a very similar process, whereby the mPFC sends requests to the hippocampus to explore possible outcomes, the value of which can then be evaluated by the ventral striatum.

Exploration has been hypothesized to be a form of information foraging and, analogously, deliberation has been hypothesized to consist of a mental exploration of one's internal schema[58]. The latter hypothesis requires that an animal has a model of the world (a mental schema) on which to search. Consistent with this hypothesis, VTE does not occur on initial exploration, but only after the rat has a schema about the task and environment. Similarly, VTE does not occur during every decision, but only during difficult decisions (that is, where options are similarly valued[67]) and before automation[14,53,54,56,65,68].

Conclusion

In his 1935 critique of Tolman's VTE hypothesis, Guthrie[152] said that: "So far as the theory is concerned the rat is left buried in thought". We now know that an internal search through a representation can derive information that is hidden within that representation, and thus that a search through a mental schema of the consequences of one's actions can derive information that can change decisions. Importantly, computation takes time. The search-and-evaluate process that underlies deliberation therefore requires time to complete. Although several questions remain (BOX 6), the term 'vicarious trial and error' makes strong assumptions about cognitive and mental states; it implies that rats are actually searching through possibilities, evaluating those possibilities and making decisions that are based on those evaluations. In other words, it implies that rats really are deliberating. The data that have come in over the past decade are clear: Muenzinger, Gentry and Tolman were completely correct — VTE is a process of vicarious (mental) trial and error (search, evaluate and test). Guthrie was also right: when rats pause at choice points, they are indeed "buried in thought".

1. Muenzinger, K. F. & Gentry, E. Tone discrimination in white rats. *J. Comp. Psychol.* **12**, 195–206 (1931).
2. Tolman, E. C. Prediction of vicarious trial and error by means of the schematic sowbug. *Psychol. Rev.* **46**, 318–336 (1939).
3. Tolman, E. C. Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208 (1948).
4. Hull, C. L. *Principles of Behavior* (Appleton-Century-Crofts, 1943).
5. Turing, A. On computable numbers, with an application to the entscheidungs problem. *Proc. Lond. Math. Soc.* **42**, 230–265 (1937).
6. Wiener, N. *Cybernetics, or Control and Communications in the Animal and the Machine* (Hermann, 1948).
7. Shannon, C. A mathematical theory of communication. *Bell System Techn. J.* **27**, 379–423, 623–656 (1948).
8. Newell, A., Shaw, J. C. & Simon, H. A. in *Proc. Int. Conf. Information Process*. [online], http://bitsavers.informatik.uni-stuttgart.de/pdf/rand/ipl/P-1584_Report_On_A_General_Problem-Solving_Program_Feb59.pdf (UNESCO,1959).
9. Simon, H. A behavioral model of rational choice. *Q. J. Econ*. **69**, 99–118 (1955).
10. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Clarendon Press, 1978).
This is a comprehensive book laying out the hypothesis that the hippocampus has a key role in Tolman's cognitive map. Also, this book explains one of the first proposed algorithmic differences between what is now called deliberative and procedural learning.
11. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* **261**, 1055–1058 (1993).
12. Zhang, K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J. Neurosci.* **16**, 2112–2126 (1996).
13. Brown, E. N., Frank, L. M., Tang, D., Quirk, M. C. & Wilson, M. A. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J. Neurosci.* **18**, 7411–7425 (1998).
14. Johnson, A. & Redish, A. D. Neural ensembles CA3 transiently encode paths forward animal decision point. *J. Neurosci.* **27**, 12176–12189 (2007).
This paper shows for the first time that hippocampal sequences encode future outcomes at choice points.
15. Foster, D. J. & Wilson, M. A. Hippocampal theta sequences. *Hippocampus* **17**, 1093–1099 (2007).
16. Pfeiffer, B. E. & Foster, D. J. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**, 74–79 (2013).
17. O'Craven, K. M. & Kanwisher, N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* **12**, 1013–1023 (2000).
18. Pearson J., Naselaris, T., Holmes, E. A. & Kosslyn S. M. Mental imagery: functional mechanisms clinical application. *Trends Cogn. Sci.* **19**, 590–602 (2015).
This is a clear review of the now established fact that imagination activates the same circuits as perception in humans, which can be used (in reference 19) to identify mental time travel in non-human animals.
19. Johnson, A., Fenton, A. A., Kentros, C. & Redish, A. D. Looking for cognition in the structure within the noise. *Trends Cogn. Sci.* **13**, 55–64 (2009).
This paper lays out the logic and mathematics allowing for identification of neural representations of cognitive events (such as mental time travel) from neural signals in non-human animals.
20. Payne, J., Bettman, J. & Johnson, E. *The Adaptive Decision Maker* (Cambridge Univ. Press, 1993).
21. Gilbert, D. T. & Wilson, T. D. Prospection: experiencing the future. *Science* **317**, 1351–1354 (2007).
22. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
23. Kurth-Nelson, Z., Bickel, W. K. & Redish, A. D. A theoretical account of cognitive effects in delay discounting. *Eur. J. Neurosci.* **35**, 1052–1064 (2012).
24. Redish, A. D. *The Mind within the Brain: How we Make Decisions and How Those Decisions Go Wrong* (Oxford Univ. Press, 2013).
This is a thorough review of the concepts and theory underlying the multiple-decision systems hypothesis.
25. Buckner, R. L. & Carroll, D. C. Self-projection and the brain. *Trends Cognitive Sci.* **11**, 49–57 (2007).
26. Schacter, D. L. & Addis, D. R. in *Predictions in the Brain: Using Our Past to Generate a Future* (ed. Bar, M.) 58–69 (Oxford Univ. Press, 2011).
27. Gilbert, D. T. & Wilson, T. D. Why the brain talks to itself: sources of error in emotional prediction. *Phil. Trans. R Soc. B* **364**, 1335–1341 (2009).
28. Phelps, E., Lempert, K. M. & Sokol-Hessner, P. Emotion and decision making: multiple modulatory circuits. *Annu. Rev. Neurosci.* **37**, 263–287 (2014).
29. Redish, A. D., Schultheiss, N. W. & Carter, E. C. The computational complexity of valuation and motivational forces in decision-making processes. *Curr. Top. Behav. Neurosci.* http://dx.doi.org/10.1007/7854_2015_375 (2015).
30. Redish, A. D., Jensen, S., Johnson, A. & Kurth-Nelson, Z. Reconciling reinforcement learning models with behavioral extinction and renewal: Implications for addiction, relapse, and problem gambling. *Psychol. Rev.* **114**, 784–805 (2007).
31. Gershman, S. J., Blei, D. & Niv, Y. Context, learning and extinction. *Psychol. Rev.* **117**, 197–209 (2010).
32. Niv, Y. *et al.* Reinforcement learning in multidimensional environments relies on attention mechanisms. *J. Neurosci.* **35**, 8145–8157 (2015).
33. Johnson, A. & Crowe, D. A. Revisiting Tolman, his theories and cognitive maps. *Cogn. Critique* **1**, 43–72 (2009).
34. Redish, A. D. *Beyond the Cognitive Map: From Place Cells to Episodic Memory* (MIT Press, 1999).
35. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).
36. Winstanley, C. A. *et al.* in *Cognitive Search: Evolution, Algorithms, and the Brain* (eds Hills, T. *et al.*) 125–156 (MIT Press, 2012).
37. McKenzie, S. *et al.* Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron* **83**, 202–215 (2014).
38. Lichtenstein, S. & Slovic, P. (eds) *The Construction of Preference* (Cambridge Univ. Press, 2006).
39. Kahneman, D. *Thinking, Fast, and Slow* (Farrar, 2011).
40. Benoit, R. G., Gilbert, S. J. & Burgess, P. W. A neural mechanism mediating the impact of episodic prospection on farsighted decisions. *J. Neurosci.* **31**, 6771–6779 (2011).

41. Peters, J. & Büchel, C. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal–mediotemporal interactions. *Neuron* **66**, 138–148 (2010).
42. Kwan, D. *et al.* Future decision-making without episodic mental time travel. *Hippocampus* **22**, 1215–1219 (2012).
43. Hassabis, D., Kumaran, D., Vann, S. D. & Maguire, E. A. Patients with hippocampal amnesia cannot imagine experiences. *Proc. Natl Acad. Sci. USA* **104**, 1726–1731 (2007).
This paper shows that the hippocampus is critical for the ability to create episodic futures in humans.
44. Lebreton, M. *et al.* A critical role for the hippocampus in the valuation of imagined outcomes. *PLoS Biol.* **11**, e1001684 (2013).
45. Wang, J. X., Cohen, N. J. & Voss, J. L. Covert rapid action-memory simulation (crams): a hypothesis of hippocampal–prefrontal interactions for adaptive behavior. *Neurobiol. Learn. Mem.* **117**, 22–33 (2015).
46. Spiers, H. J. & Gilbert, S. J. Solving the detour problem in navigation: a model of prefrontal and hippocampal interactions. *Front. Hum. Neurosci.* **9**, 125 (2015).
47. Hikosaka, O. *et al.* Parallel neural networks for learning sequential procedures. *Trends Neurosci.* **22**, 464–471 (1999).
48. Lee, S. W., Shimoko, S. & O'Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **81**, 687–699 (2014).
49. Dezfouli, A. & Balleine, B. Habits, action sequences and reinforcement learning. *Eur. J. Neurosci.* **35**, 1036–1051 (2012).
50. Johnson, A., van der Meer, M. A. & Redish, A. D. Integrating hippocampus and striatum in decision-making. *Curr. Opin. Neurobiol.* **17**, 692–697 (2007).
51. Gardner, R. S. *et al.* A secondary working memory challenge preserves primary place strategies despite overtraining. *Learn. Mem.* **20**, 648–656 (2013).
On the classic Tolman–Hull plus maze, VTE arises with deliberative choices.
52. Schmidt, B. J., Papale, A. E., Redish, A. D. & Markus, E. J. Conflict between place and response navigation strategies: effects on vicarious trial and error (VTE) behaviors. *Learn. Mem.* **20**, 130–138 (2013).
53. Smith, K. S. & Graybiel, A. M. A dual operator view of habitual behavior reflecting cortical striatal dynamics. *Neuron* **79**, 361–374 (2013).
This paper shows that VTE is negatively related to dorsolateral striatal representations that reflect procedural strategies (task bracketing).
54. Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C. & Redish, A. D. Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron* **67**, 25–32 (2010).
This paper shows that hippocampal representations encode future options during VTE, ventral striatal representations encode potential rewards during VTE, and dorsolateral striatum does neither, instead slowly developing situation-action representations in line with the automation of behaviour.
55. Thorn, C. A., Atallah, H., Howe, M. & Graybiel, A. M. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* **66**, 781–795 (2010).
56. Regier, P. S., Amemiya, S. & Redish, A. D. Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task. *J. Neurophysiol.* **114**, 1399–1416 (2015).
57. Niv, Y., Joel, D. & Dayan, P. A normative perspective on motivation. *Trends Cogn. Sci.* **10**, 375–381 (2006).
This paper gives an explication of the difference between search-and-evaluate and cached-action-chain decision systems.
58. Johnson, A., Varberg, Z., Benhardus, J., Maahs, A. &Schrater, P. The hippocampus and exploration: dynamically evolving behavior and neural representations. *Front. Hum. Neurosci.* **6**, 216 (2012).
59. Balleine, B. W. & Dickinson, A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407–419 (1998).
60. Liljeholm, M., Tricomi, E., O'Doherty, J. P. & Balleine, B. W. Neural correlates of instrumental contingency learning: differential effects of actionreward conjunction and disjunction. *J. Neurosci.* **31**, 2474–2480 (2011).
61. Krajbich, I. & Rangel, A. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl Acad. Sci. USA* **108**, 13852–13857 (2011).
Investigating human saccade–fixate–saccade sequences, this paper suggests a decision model and shows that it is consistent with the behavioural selections made by subjects.
62. Packard, M. G. & McGaugh, J. L. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol. Learn. Mem.* **65**, 65–72 (1996).
63. Yin, H. H., Knowlton, B. & Balleine, B. W. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* **19**, 181–189 (2004).
64. Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S. & Redish, A. D. Hippocampal replay is not a simple function of experience. *Neuron* **65**, 695–705 (2010).
65. Blumenthal, A., Steiner, A., Seeland, K. D. & Redish, A. D. Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task. *Neurobiol. Learn. Mem.* **95**, 376–384 (2011).
66. Steiner, A. & Redish, A. D. The road not taken: neural correlates of decision making in orbitofrontal cortex. *Front. Decision Neurosci.* **6**, 131 (2012).
67. Steiner, A. & Redish, A. D. Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat. Neurosci.* **17**, 995–1002 (2014).
68. Papale, A., Stott, J. J., Powell, N. J., Regier, P. S. & Redish, A. D. Interactions between deliberation and delay-discounting in rats. *Cogn. Affect. Behav. Neurosci.* **12**, 513–526 (2012).
On the spatial delay-discounting task, VTE occurs during the titration phase, when the rat is making flexible choices, not during the exploitation phase, when the rat has automated its behaviour, even though the values of the two options are equal in the exploitation phase.
69. Atance, C. M. & O'Neill, D. K. Episodic future thinking. *Trends Cogn. Sci.* **5**, 533–539 (2001).
70. Seidenbecher, T., Laxmi, T. R., Stork, O. & Pape, H. C. Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. *Science* **301**, 846–850 (2003).
71. Schmidt, R. *et al.* Single-trial phase precession in the hippocampus. *J. Neurosci.* **29**, 13232–13241 (2009).
72. Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S. & Redish, A. D. Segmentation of spatial experience by hippocampal θ sequences. *Nat. Neurosci.* **15**, 1032–1039 (2012).
73. Hassabis, D. & Maguire, E. A. in *Predictions in the Brain: Using our Past to Generate a Future* (ed. Bar, M.) 70–82 (Oxford Univ. Press, 2011).
74. Howard, L. R. *et al.* The hippocampus and entorhinal cortex encode the path and Euclidean distances to goals during navigation. *Curr. Biol.* **24**, 1331–1340 (2014).
75. Wikenheiser, A. M. & Redish, A. D. Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* **18**, 289–294 (2015).
76. Royer, S., Sirota, A., Patel, J. & Buzsaki, G. Distinct representations and theta dynamics in dorsal and ventral hippocampus. *J. Neurosci.* **30**, 1777–1787 (2010).
77. Stensola, H. *et al.* The entorhinal grid map is discretized. *Nature* **492**, 72–78 (2012).
78. Dragoi, G. & Buzsaki, G. Temporal encoding of place sequences by hippocampal cell assemblies. *Neuron* **50**, 145–157 (2006).
79. Lisman, J. & Redish, A. D. Prediction, sequences and the hippocampus. *Phil. Trans. R. Soc. B* **364**, 1193–1201 (2009).
80. Buzsaki, G. Hippocampal sharp wave-ripple: a cognitive biomarker for episodic memory and planning. *Hippocampus* **25**, 1073–1188 (2015).
81. Diba, K. & Buzsaki, G. Forward and reverse hippocampal place-cell sequences during ripples. *Nat. Neurosci.* **10**, 1241–1242 (2007).
82. Yamamoto, J., Suh, J., Takeuchi, D. & Tonegawa, S. Successful execution of working memory linked to synchronized high-frequency gamma oscillations. *Cell* **157**, 845–857 (2014).
83. Schmitzer-Torbert, N. C. & Redish, A. D. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple-T task. *J. Neurophysiol.* **91**, 2259–2272 (2004).
84. Tse, D. *et al.* Schemas and memory consolidation. *Science* **316**, 76–82 (2007).
85. Hu, D. & Amsel, A. A simple test of the vicarious trial-and-error hypothesis of hippocampal function. *Proc. Natl Acad. Sci. USA* **92**, 5506–5509 (1995).
86. Bett, D. *et al.* The neural substrates of deliberative decision making: contrasting effects of hippocampus lesions on performance and vicarious trial-and-error behavior in a spatial memory task and a visual discrimination task. *Front. Behav. Neurosci.* **6**, 70 (2012).
87. Bett, D., Murdoch, L. H., Wood, E. R. & Dudchenko, P. A. Hippocampus, delay discounting, and vicarious trial-and-error. *Hippocampus* **25**, 643–654 (2015).
88. Barnes, C. A., Nadel, L. & Honig, W. K. Spatial memory deficit in senescent rats. *Can. J. Psychol.* **34**, 29–39 (1980).
89. Breton, Y. A., Seeland K. D. & Redish, A. D. Aging impairs deliberation and behavioral flexibility in inter-temporal choice. *Front. Aging Neurosci.* **7**, 41 (2015).
90. Robbe, D. *et al.* Cannabinoids reveal importance of spike timing coordination in hippocampal function. *Nat. Neurosci.* **9**, 1526–1533 (2006).
91. Robbe, D. & Buzsaki, G. Alteration of theta timescale dynamics of hippocampal place cells by a cannabinoid is associated with memory impairment. *J. Neurosci.* **29**, 12597–12605 (2009).
92. Papale, A. E. *Hippocampal Representations on the Spatial Delay Discounting Task.* Thesis, Univ. Minnesota (2015).
93. Amemiya, S., Noji, T., Kubota, N., Nishijima, T. & Kita, I. Noradrenergic modulation of vicarious trial-and-error behavior during a spatial decision-making task in rats. *Neuroscience* **265**, 291–301 (2014).
94. Amemiya, S. & Redish, A. D. Manipulating decisiveness in decision making: effects of clonidine on hippocampal search strategies. *J. Neurosci.* **36**, 814–827 (2016).
95. Seifert, W. *Neurobiology of the Hippocampus* (Academic Press, 1983).
96. Teyler, T. J. and DiScenna, P. The hippocampal memory indexing theory. *Behav. Neurosci.* **100**, 147–154 (1986).
97. Nadel, L., Samsonovich, A., Ryan, L. & Moscovitch, M. Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* **10**, 352–368 (2000).
98. Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D. & Daw, N. D. Model-based choices involve prospective neural activity. *Nat. Neurosci.* **18**, 767–772 (2015).
In humans, prospection entails the representation of specific future outcomes.
99. Killcross, S. & Coutureau, E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* **13**, 400–408 (2003).
100. Sharpe, M. J. & Killcross, S. The prelimbic cortex directs attention toward predictive cues during fear learning. *Learn. Mem.* **22**, 289–293 (2015).
101. Spellman, T. *et al.* Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature* **522**, 309–314 (2015).
102. Ito, H. T., Zhang, S. J., Witter, M. P., Moser, E. I. & Moser, M. B. A prefrontal–thalamo–hippocampal circuit for goal-directed spatial navigation. *Nature* **522**, 50–55 (2015).
103. Kolb, B. in *The Cerebral Cortex of the Rat* (eds Kolb & Tees R. C.) 437–458 (MIT Press, 1990).
104. Dalley, J. W., Cardinal, R. N. & Robbins, T. W. Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neurosci. Biobehav. Rev.* **28**, 771–784 (2004).
105. Wallis, J. D. Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat. Neurosci.* **15**, 13–19 (2012).
106. Jung, M. W., Qin, Y., McNaughton, B. L. & Barnes, C. A. Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. *Cereb. Cortex* **8**, 437–450 (1998).
107. Powell, N. J. & Redish, A. D. Complex neural codes in rat prelimbic cortex are stable across days on a spatial decision task. *Front. Behav. Neurosci.* **8**, (2014).
108. Jones, M. W. & Wilson, M. A. Theta rhythms coordinate hippocampal–prefrontal interactions in a spatial memory task. *PLoS Biol.* **3**, e402 (2005).
109. Hyman, J. M., Zilli, E. A., Paley, A. M. & Hasselmo, M. E. Working memory performance correlates with prefrontal-hippocampal theta interactions but not with prefrontal neuron firing rates. *Front. Integr.Neurosci.* **4**, 2 (2010).

110. Ragozzino, M. E., Detrick, S. & Kesner, R. P. Involvement of the prelimbic–infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *J. Neurosci.* **19**, 4585–4594 (1999).
111. Rich, E. L. & Shapiro, M. L. Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *J. Neurosci.* **27**, 4747–4755 (2007).
112. Benchenane, K. *et al.* Coherent theta oscillations and reorganization of spike timing in the hippocampal–prefrontal network upon learning. *Neuron* **66**, 921–936 (2010).
113. O'Neill, P. K., Gordon, J. A. & Sigurdsson, T. Theta oscillations in the medial prefrontal cortex are modulated by spatial working memory and synchronize with the hippocampus through its ventral subregion. *J. Neurosci.* **33**, 14211–14224 (2013).
114. Engel, A. K., Fries, P. & Singer, W. Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* **2**, 704–716 (2001).
115. Kucewicz, M. T., Tricklebank, M. D., Bogacz, R. & Jones, M. W. Dysfunctional prefrontal cortical network activity and interactions following cannabinoid receptor activation. *J. Neurosci.* **31**, 15560–15568 (2011).
116. Mogenson, G. J., Jones, D. L. & Yim, C. Y. From motivation to action: functional interface between the limbic system and the motor system. *Prog. Neurobiol.* **14**, 69–97 (1980).
117. McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y. & Schoenbaum, G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J. Neurosci.* **31**, 2700–2705 (2011).
118. Jones, J. L. *et al.* Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* **338**, 953–956 (2012).
119. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
120. Strait, C. E., Blanchard, T. C. & Hayden, B. Y. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* **82**, 1357–1366 (2014).
121. Strait, C. E., Sleezer, B. J. & Hayden, B. Y. Signatures of value comparison in ventral striatum neurons. *PLoS Biol.* **13**, e1002173 (2015).
122. Bechara, A., Tranel, D. & Damasio, H. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* **123**, 2189–2202 (2000).
123. Zald, D. H. & Rauch, S. L. (eds) *The Orbitofrontal Cortex* (Oxford Univ. Press, 2008).
124. Fellows, L. K. Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. *Brain* **129**, 944–952 (2006).
125. Bray, S., Shimojo, S. & O'Doherty, J. P. Human medial orbitofrontal cortex is recruited during experience of imagined and real rewards. *J. Neurophysiol.* **103**, 2506–2512 (2010).
126. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
127. Nicola, S. M., Yun, I. A., Wakabayashi, K. T. & Fields, H. L. Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task. *J. Neurophysiol.* **91**, 1840–1865 (2004).
128. Roitman, M. F., Wheeler, R. A. & Carelli, R. M. 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 (2005).
129. Stott, J. J. & Redish, A. D. A functional difference in information processing between orbitofrontal cortex ventral striatum during decision-making behaviour. *Phil. Trans. R. Soc. B* **369**, 20130472 (2014). **In simultaneous recordings from the ventral striatum and the orbitofrontal cortex, ventral striatal representations show covert reward-related information before the turn-around during VTE, but orbitofrontal representations only show covert reward-related information after the turn-around point.**
130. German, P. W. & Fields, H. L. Rat nucleus accumbens neurons persistently encode locations associated with morphine reward. *J. Neurophysiol.* **97**, 2094–2106 (2007).
131. Moorman, D. E. & Aston-Jones, G. Orbitofrontal cortical neurons encode expectation-driving initiation of reward-seeking. *J. Neurosci.* **34**, 10234–10246 (2014).
132. van der Meer, M. A. A. & Redish, A. D. Covert expectation-of-reward in rat ventral striatum at decision points. *Front. Integr. Neurosci.* **3**, 1–15 (2009).
133. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279 (2014).
134. Fenton, A. A. *et al.* Attention-like modulation of hippocampus place cell discharge. *J. Neurosci.* **30**, 4613–4625 (2010).
135. Wikenheiser, A. & Redish, A. D. Changes in reward contingency modulate the trial to trial variability of hippocampal place cells. *J. Neurophysiol.* **106**, 589–598 (2011).
136. Berke, J. D. & Eichenbaum, H. Striatal versus hippocampal representations during win-stay maze performance. *J. Neurophysiol.* **101**, 1575–1587 (2009).
137. Schmitzer-Torbert, N. C. & Redish, A. D. Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. *Neuroscience* **153**, 349–360 (2008).
138. Jog, M. S., Kubota, Y., Connolly, C. I., Hillegaart, V. & Graybiel, A. M. Building neural representations of habits. *Science* **286**, 1746–1749 (1999).
139. Jin, X. & Costa, R. M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457–462 (2010).
140. Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z. & Graybiel, A. M. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* **437**, 1158–1161 (2005).
141. Barnes, T. D. *et al.* Advance cueing produces enhanced action-boundary patterns of spike activity in the sensorimotor striatum. *J. Neurophysiol.* **105**, 1861–1878 (2011).
142. Kubota, Y. *et al.* Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. *J. Neurophysiol.* **102**, 2142–2160 (2009).
143. Corbit, L. H. & Balleine, B. W. The role of the hippocampus in instrumental conditioning. *J. Neurosci.* **20**, 4233–4239 (2000).
144. Ragozzino, M. E., Ragozzino, K. E., Mizumori, S. J. Y. & Kesner, R. P. The role of the dorsomedial striatum in behavioral flexibility for repsonse and visual cue discrimination learning. *Behav. Neurosci.* **116**, 105–115 (2002).
145. Ostlund, S. B., Winterbauer, N. E. & Balleine, B. W. Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. *J. Neurosci.* **29**, 8280–8287 (2009).
146. Corbit, L. H. & Janak, P. H. Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *Eur. J. Neurosci.* **31**, 1312–1321 (2010).
147. Bower, G. H. in *Studies in Mathematical Learning Theory* (eds Bush, R. R. & Estes, W. K.) 109–124 (Stanford Univ. Press, 1959).
148. Whishaw, I. Q. & Kolb, B. *The Behavior of the Laboratory Rat: A Handbook with Tests* (Oxford Univ. Press, 2004).
149. Monaco, J. D., Rao, G., Roth, E. D. & Knierim, J. J. Attentive scanning behavior drives one-trial potentiation of hippocampal place fields. *Nat. Neurosci.* **17**, 725–731 (2014).
150. Baddeley, A. *Working Memory, Thought and Action* (Oxford Univ. Press, 2007).
151. Fuster, J. *The Prefrontal Cortex* (Academic Press, 2008).
152. Guthrie, E. R. *The Psychology of Learning* (Harpers, 1935).
153. Rodgers, R. J., Cao, B. J., Dalvi, A. & Holmes, A. Animals models of anxiety: an ethological perspective. *Brazil. J. Med. Biol. Res.* **30**, 289–304 (1997).
154. Jackson, J. C. & Redish, A. D. Detecting dynamical changes within a simulated neural ensemble using a measure of representational quality. *Network* **14**, 629–645 (2003).
155. Adhikari, A. Distributed circuits underlying anxiety. *Front. Behav. Neurosci.* **8**, 112 (2014).
156. Allsop, S. A., Vander Weele, C. M., Wichmann, R. & Tye, K. M. Optogenetic insights on the relationship between anxiety-related behaviors and social deficits. *Front. Behav. Neurosci.* **8**, 241 (2014).
157. Adhikari, A., Topiwala, M. A. & Gordon, J. A. Synchronized activity between the ventral hippocampus and the medial prefrontal cortex during anxiety. *Neuron* **28**, 257–269 (2010).
158. Adhikari, A., Topwala, M. & Gordon, J. A. Single units in the medial prefrontal cortex with anxiety-related firing patterns are preferentially influenced by ventral hippocampal activity. *Neuron* **71**, 898–910 (2011).
159. Grewal, S. S., Shepherd, J. K., Bill, D. J., Fletcher, A. & Dourish, C. T. Behavioural and pharmacological characterisation of the canopy stretched attend posture test as a model of anxiety in mice and rats. *Psychopharmacology* **133**, 29–38 (1997).
160. Choi, J. S. & Kim, J. J. Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proc. Natl Acad. Sci. USA* **107**, 21773–21777 (2010).
161. Amir, A., Lee, S. C., Headley, D. B., Herzallah, M. M. & Pare, D. Amygdala signaling during foraging in a hazardous environment. *J. Neurosci.* **35**, 12994–13005 (2015).
162. Voss, J. L. *et al.* Spontaneous revisitation during visual exploration as a link among strategic behavior, learning, and the hippocampus. *Proc. Natl Acad. Sci. USA* **108**, E402–E409 (2011).
163. Reutskaja, E., Nagel, R., Camerer, C. F. & Rangel, A. Search dynamics in consumer choice under time pressure: an eye-tracking study. *Am. Econom. Rev.* **101**, 900–926 (2011).
164. Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **12**, 97–136 (1980).
165. Kruse, J. M., Overmier, J. B., Konz, W. A. & Rokke, E. Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. *Learn. Motiv.* **14**, 165–181 (1983).
166. Balleine, B. W. & Ostlund, S. B. Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann. NY Acad. Sci.* **1104**, 147–171 (2007).
167. Talmi, D., Seymour, B., Dayan, P. & Dolan, R. J. Human Pavlovian-instrumental transfer. *J. Neurosci.* **28**, 360–368 (2008).
168. Bray, S., Rangel, A., Shimojo, S., Balleine, B. & O'Doherty, J. P. The neural mechanisms underlying the influence of Pavlovian cues on human decision making. *J. Neurosci.* **28**, 5861–5866 (2008).
169. Damasio, A. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain* (Mariner, 2003).
170. Balleine, B. W. in *The Behavior of the Laboratory Rat: A Handbook with Tests* (eds Whishaw, I. Q. & Kolb, B.) 436–446 (Oxford Univ. Press, 2004).
171. Gold, J. I. & Shadlen, M. N. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* **36**, 299–308 (2002).
172. Hanks, T. D. *et al.* Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015).

Acknowledgements
The author thanks all of the students and colleagues who have worked with him over the years looking at different aspects of VTE, both for helpful discussions and for the experimental and theoretical work on these questions. The author also thanks A. Johnson, K. Smith, J. Stott, S. Amemiya and Y. Breton for comments on drafts of this manuscript, and J. Voss for helpful discussions. Time to work on this manuscript was funded by MH080318 and DA030672.

Competing interests statement
The author declares no competing interests.