



# Decoding cognition from spontaneous neural activity

Yunzhe Liu<sup>1,2,3</sup>✉, Matthew M. Nour<sup>3,4</sup>, Nicolas W. Schuck<sup>3,5</sup>, Timothy E. J. Behrens<sup>4,6</sup> and Raymond J. Dolan<sup>1,3,4,7</sup>

**Abstract** | In human neuroscience, studies of cognition are rarely grounded in non-task-evoked, ‘spontaneous’ neural activity. Indeed, studies of spontaneous activity tend to focus predominantly on intrinsic neural patterns (for example, resting-state networks). Taking a ‘representation-rich’ approach bridges the gap between cognition and resting-state communities: this approach relies on decoding task-related representations from spontaneous neural activity, allowing quantification of the representational content and rich dynamics of such activity. For example, if we know the neural representation of an episodic memory, we can decode its subsequent replay during rest. We argue that such an approach advances cognitive research beyond a focus on immediate task demand and provides insight into the functional relevance of the intrinsic neural pattern (for example, the default mode network). This in turn enables a greater integration between human and animal neuroscience, facilitating experimental testing of theoretical accounts of intrinsic activity, and opening new avenues of research in psychiatry.

An overarching goal of neuroscience is to understand the relationship between cognition and underlying neural activity. In humans, much progress in this direction has been driven by task-based neuroimaging studies. However, it is salutary to note that a large proportion of variance in neural activity is off-task<sup>1,2</sup>. We use the term ‘spontaneous’ to refer to intrinsic activities that are not a mere response to external events. We review past work and an emerging new paradigm for studying cognition in the spontaneous neural activity.

## A dichotomy in human neuroscience

In cognitive neuroscience, a standard approach is to temporally align neural activity to specific task events (for example, presentation of a visual stimulus, such as a ‘house’), and to localize the corresponding neural response in the brain (FIG. 1a). In this way, neuroscientists have provided a rich characterization of cognition in the context of task-evoked processing.

A major difficulty in studying spontaneous neural activity, as compared with task-evoked activity, is that researchers do not have direct access to either the identity or the timing of putative states hypothesized to drive this neural activity (hence the term ‘spontaneous’). This makes attribution to a causal cognitive process less straightforward. Broadly speaking, we can conceptualize two broad approaches, spanning two ends of a continuum (FIG. 1b).

On one hand, neuroscientists adopt a ‘data-driven’ approach (FIG. 1b, right). Here, instead of a concern

with changes in neural activity ‘triggered’ by external events, the focus is on intrinsic physiological features, such as functional connectivity, power or phase coupling of frequency-specific oscillations (for example, alpha rhythm)<sup>3</sup>. The putative cognitive role of these features is sometimes inferred on the basis of their relationship to behavioural or psychological measures<sup>3,4</sup>, or by evaluating an anatomical overlap with task-evoked activity patterns reported in other studies<sup>5,6</sup>. This approach aims to find a correspondence between intrinsic and task-evoked neural patterns (sometimes termed the brain’s ‘functional architecture’)<sup>7</sup>. Nevertheless, such data-driven approaches are not as well positioned to provide a cognitive grounding because a direct relationship to task events cannot be assumed (unlike in task-evoked studies, where neural activity and task events can be temporally aligned).

At the other end of the spectrum, researchers have studied spontaneous activity using approaches that bear similarity to the analysis of task-evoked activity, which we term ‘intervention-based approaches’. These approaches introduce an external intervention known to evoke an associated event within an otherwise task-free session (FIG. 1b, left). Examples here include using targeted memory reactivation<sup>8,9</sup> in sleep studies<sup>10,11</sup>, tracking spontaneous neural activity during prestimulus time<sup>12</sup> or presenting stimuli at the peak (strongest time) or trough (weakest time) of neural activity of interest (for example, dopaminergic midbrain)<sup>13</sup>. In this way, spontaneous activity can be analysed by alignment to

<sup>1</sup>State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China.

<sup>2</sup>Chinese Institute for Brain Research, Beijing, China.

<sup>3</sup>Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London, UK.

<sup>4</sup>Wellcome Centre for Human Neuroimaging, University College London, London, UK.

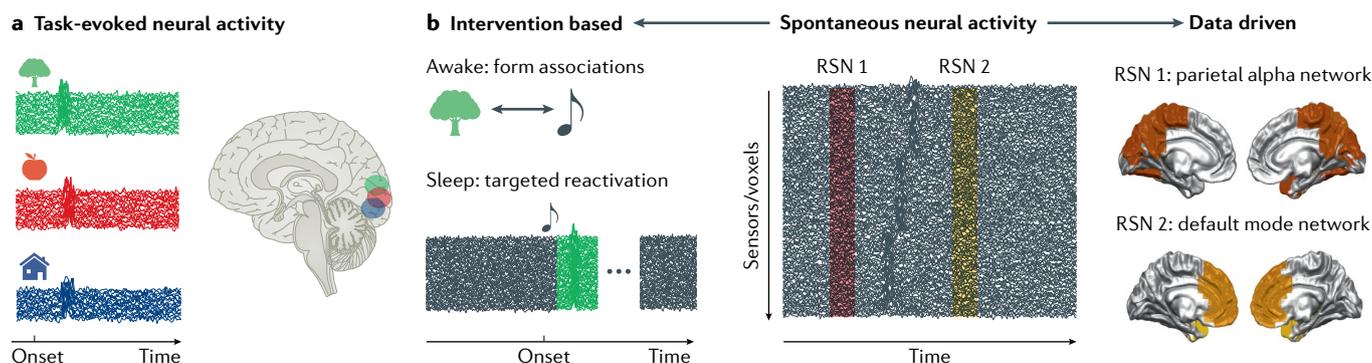
<sup>5</sup>Max Planck Research Group Neurocode, Max Planck Institute for Human Development, Berlin, Germany.

<sup>6</sup>Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK.

<sup>7</sup>Department of Psychiatry, Universitätsmedizin Berlin, Berlin, Germany.

✉e-mail: [yunzhe.liu@bnu.edu.cn](mailto:yunzhe.liu@bnu.edu.cn)

<https://doi.org/10.1038/s41583-022-00570-z>



**Fig. 1 | A (relative) dichotomy in human neuroscience. a** | Cognitive processes are typically studied by aligning neural activity to the onset of perturbing stimuli, with the aim of finding underlying neural correlates. For example, with use of functional MRI, evoked neural activity in response to different stimuli (or task demands) is mapped in the brain, here illustrated by responses to a house (blue), an apple (red) and a tree (green), respectively. **b** | Studies of spontaneous neural activity can be rendered more akin to task-evoked studies by introducing an external intervention. For example, in a targeted memory reactivation approach, a tone associated with a specific stimulus is used to evoke task-related processing during sleep

(left; green indicates tree-related processing, elicited by its paired tone). Alternatively, a data-driven approach can be used to characterize the canonical functional connectivity patterns during rest (right). Two examples of resting-state networks (RSNs; parietal alpha network (red) and default mode network (yellow)). The colour reflects the broadband power, where a brighter colour indicates higher power. Although there may be task-related reactivations during rest (for example, transient synchronous activity bumps in between RSN 1 and RSN 2 epochs, shown in darker patch), their functional relevance is not accessible to methods used in standard resting state studies. Part **b** adapted with permission from REF.<sup>15</sup>, Elsevier.

- Task**  
An experiment designed to manipulate an assumed cognitive process.
- Off-task**  
Period without explicit task demand (for example, during rest).
- Functional connectivity**  
Temporal dependency of neuronal activation (for example, correlation) between anatomically separated brain regions.
- Power**  
The strength of a signal in a given frequency band.
- Decoding**  
Reading out task-related information from neural activity.
- Pairwise multivoxel correlation**  
Correlation between all pairs of voxels of interest using the entire time course of the functional MRI signal.
- Representational similarity analysis**  
Measure of the similarity of neural activity among different conditions.

the onset of such events, with precise timing akin to task-evoked studies. Unlike a data-driven approach, this intervention-based approach enables direct inference with respect to the underlying cognitive process. This benefit, however, comes at the expense of interrupting internal computations that are ‘spontaneous’ in nature, such as imagination or mind wandering — processes that are by definition not tied to immediate task demand<sup>14</sup>. Thus, cognitive neuroscientists were faced with a choice: to study unperturbed spontaneous neural activity, albeit with a restricted window to its functional relevance, or instead to have more explicit control of the cognitive process, but at the expense of disrupting spontaneity.

In recent years, an emerging approach has endeavoured to combine the best of both techniques outlined above. This approach relies on exploiting the representational content of neural activity and is predicated on an assumption that the same neural representations of task events (for example, ‘house’ or ‘apple’) are active both on-task and off-task. Such representations can first be derived from task-evoked neural activity and their reactivations can subsequently be obtained through decoding. In essence, this approach probes the task-relevant content of intrinsic neural activity, going beyond a characterization from the data-driven approach<sup>15</sup>. In so doing, it provides information regarding both what and when representation has been activated, absent from external intervention. We refer to this line of research as ‘representation-rich’.

Task-related representations can be obtained in multiple ways. One approach is to rely on the neural pattern similarity of task events between on-task and off-task period<sup>16</sup> or different brain regions<sup>17</sup>. For example, to look for features of memory consolidation, Tambini and Davachi<sup>18</sup> compared the pairwise multivoxel correlation structure between stimuli at encoding and post-encoding rest versus that at pre-encoding rest, and found increased

hippocampal pattern similarity attributable to learning. This correlation-based approach has similarities to representational similarity analysis, which is widely used to study task-evoked neural activity<sup>19,20</sup>, and most often for localizing where (in the brain) a pattern emerges.

A representation-rich approach to spontaneous neural activity aims to uncover the temporal structure of task-related representations (for example, how their temporal dynamics unfold)<sup>21–25</sup>. This is typically implemented using a decoding-based method, transforming spontaneous neural activity into a time series of task-related reactivations<sup>23–25</sup>. This line of research in human neuroimaging has an interesting parallel in animal work. For example, in the rodent hippocampus, researchers have identified pyramidal cells that encode spatial locations during active navigation, known as place cells<sup>26</sup>, and also observed these same cells fire spontaneously in an organized sequence during rest. This firing recapitulates past or potential future trajectories and is referred to as ‘hippocampal replay’<sup>17,27–40</sup>. The ability to readout reactivation of specific locations during rest allows researchers to go beyond a mere characterization of neurophysiological features — for example, sharp wave ripples (SWRs)<sup>41</sup> — enabling a probing of the representational content of neural activity, particularly with respect to task variables. This feature has allowed studies of hippocampal replay to forge a link between cognition and physiology<sup>42</sup>, and in so doing, shed light on a range of cognitive functions subserved by spontaneous neural activity, including memory, learning and decision-making<sup>36,43,44</sup> (BOX 1).

Recent technical advances for characterizing task-related reactivations in human neuroimaging have inspired a series of studies investigating ‘human replay’ (FIG. 2a). These address complex forms of non-spatial cognition<sup>23,25,45</sup>, especially those informed by reinforcement learning (RL) models<sup>46–48</sup> (BOX 2). They also provide a unique opportunity to link resting dynamics

## Multivoxel patterns

Neural activity profile of multiple voxels in the brain.

## Transition matrix

A matrix that stores the probability of transition from state  $s$  to state  $s'$ .

## Regressors

Independent variables in a regression model.

of whole-brain connectivity (for example, the default mode network (DMN)) to spontaneous task-related reactivation (for example, replay)<sup>15</sup>.

In this Review, we first discuss these technical advances and then consider studies that have exploited this approach. Our aim is to demonstrate the exciting prospects afforded by a representation-rich approach in bridging task-based and spontaneous brain activity.

## Measuring spontaneous reactivations

Studies of neural dynamics in the human brain non-invasively typically involve monitoring by electroencephalography (EEG) or magnetoencephalography (MEG), performed over a several minutes or for up to hours. A 'decoding-based' approach to characterizing its task-related information involves two stages (FIG. 2b). The first identifies a mapping between neural activity and a task variable of interest. In the simplest example, these variables are discrete sensory stimuli (for example, apple or house). The mapping can be obtained by training either a discriminative model or a generative model linking object labels to their associated neural responses (for example, multivoxel patterns in the case of functional MRI (fMRI) or multichannel patterns in the case of MEG or EEG)<sup>49</sup>. Typically, the training data are obtained from an incidental 'localizer' task, in which multiple examples of task events and associated evoked neural activity are collected. This stage is similar to multivoxel pattern analysis<sup>50–52</sup> (FIG. 2b, left).

The second stage applies these trained models to neural activity obtained in other time periods of interest, which might constitute off-task rest or on-task epochs, and yields a time course of spontaneous task-related

reactivations<sup>45,53</sup> (FIG. 2b, right). In this way, researchers can access the representational content of such activity, thereby allowing comparison between the observed temporal structure of neural reactivations (for example, whether reactivation of 'house' reliably precedes that of 'apple') and patterns derived from a formal hypothesis (for example, the transition matrix in an RL-based model)<sup>25</sup>.

Note, the primary difference between an analysis of task-evoked neural responses (for example, representational similarity analysis) and an approach that focuses on spontaneous neural activity is that the former tracks neural representations by reference to the explicit timing of task events. By contrast, the temporal characterization of task-related reactivations is itself a primary research question in the analysis of spontaneous neural activity, made tractable by an assumed overlap between evoked and reactivated task representations. This assumption raises several methodological considerations.

**Methodological considerations.** The first issue relates to the fact that spontaneous neural reactivation of a task event is likely to be less pronounced than its evoked response (not least because the very neural decoders used to identify reactivations in spontaneous activity are trained using evoked neural responses). This can be considered as providing an increased risk of false negatives when one is quantifying spontaneous reactivations in off-task neural activity (type II error). A second consideration is that without 'ground truth' information about the identity and timing of task-related reactivations, the statistical inference procedure must also protect against false positives (that is, type I error), for example those arising due to nonspecific neural dynamics, such as autocorrelations<sup>53,54</sup>. Recent methodological work suggests that linear modelling with careful control of confounding regressors (for example, reactivations of other states)<sup>53</sup> and appropriate permutation-based statistical inference procedures are sufficient to deal with these concerns in human MEG or EEG, as well as rodent electrophysiology data<sup>53</sup> (FIG. 2c).

Other considerations relate to data modality-specific features. For example, in fMRI, we rely on the blood oxygenation level-dependent (BOLD) signal, which has a temporal resolution on the order of seconds, one that is almost certainly too slow to capture complete neural reactivation patterns associated with neural replay (on the order of milliseconds). Nevertheless, recent work demonstrates that such fast sequences can be detected using an fMRI decoding approach that assumes reactivation patterns cause systematic neural patterns of overlap in the delayed BOLD responses<sup>45</sup> (FIG. 2d).

## Spontaneous task-related reactivations

Investigation of spontaneous task-related reactivations in humans can be thought of as falling into two broad categories: (1) reactivation during rest (for example, off-task reactivation)<sup>15,23,25</sup>; (2) reactivation during task performance (for example, during intervals between sequential trials within a task)<sup>21,22,55–61</sup>. Spontaneous neural activity measured during off-task periods is not necessarily free from task-related influences. It is likely

### Box 1 | Neural sequences in rodents

Broadly speaking, there are two types of neural sequences that are the subject of intense investigations. One relates to sharp wave ripple (SWR) sequences, and the other relates to theta sequences<sup>104,147</sup>.

SWR sequences refer to sequences of hippocampal cell firing embedded within SWR epochs (~150-Hz local field potential oscillatory bursts). These events typically happen during rest<sup>27,31,33</sup>, but also happen during pauses in a behavioural task<sup>148</sup>. These are commonly referred to as 'replay'. The direction of replay is normally defined with respect to actual experience: with forward replay, the order of the pattern activity occur in the same order as in the actual experience, and with backward replay the order is the reverse of that in the actual experience. Both forward and backward replay are reported in the rodent literature, and have been shown to be modulated by task demands<sup>148</sup>. Although the exact function of the replay direction is still unknown, forward replay has been more associated with planning<sup>78</sup> (but see REF.<sup>149</sup>), and backward replay is more associated with learning — for example, propagating prediction error from the reward site<sup>30,37</sup>. Accumulating evidence suggests that SWR sequences (or replay) are important for mental functions as diverse as memory, learning and decision-making<sup>43,44</sup>.

Theta sequences are sequences of hippocampal place cells firing within a single theta cycle (~100–170-ms), generally proceeding from the location of the animal forward towards potential goals. Key here is the observation that during movement (for example, running through a linear track) and pausing (for example, at the decision point), place cell firing is organized within an oscillatory process reflected in a hippocampal theta rhythm (6–10 Hz). Theta sequences are dominantly forward although reverse theta sequences are also reported<sup>103</sup>. Neurophysiologically, theta sequences are typically associated with theta phase precession (see REF.<sup>150</sup> for independent theta phase coding), in which the firings of a particular place cell are embedded within progressively earlier phase of the theta cycle (phase precession) as the animal traverses its place field<sup>151</sup>. Functionally, theta sequences may reflect planning process<sup>105,152,153</sup>, memory formation<sup>104,147</sup> or multiple prospective futures in alteration<sup>153,154</sup>.

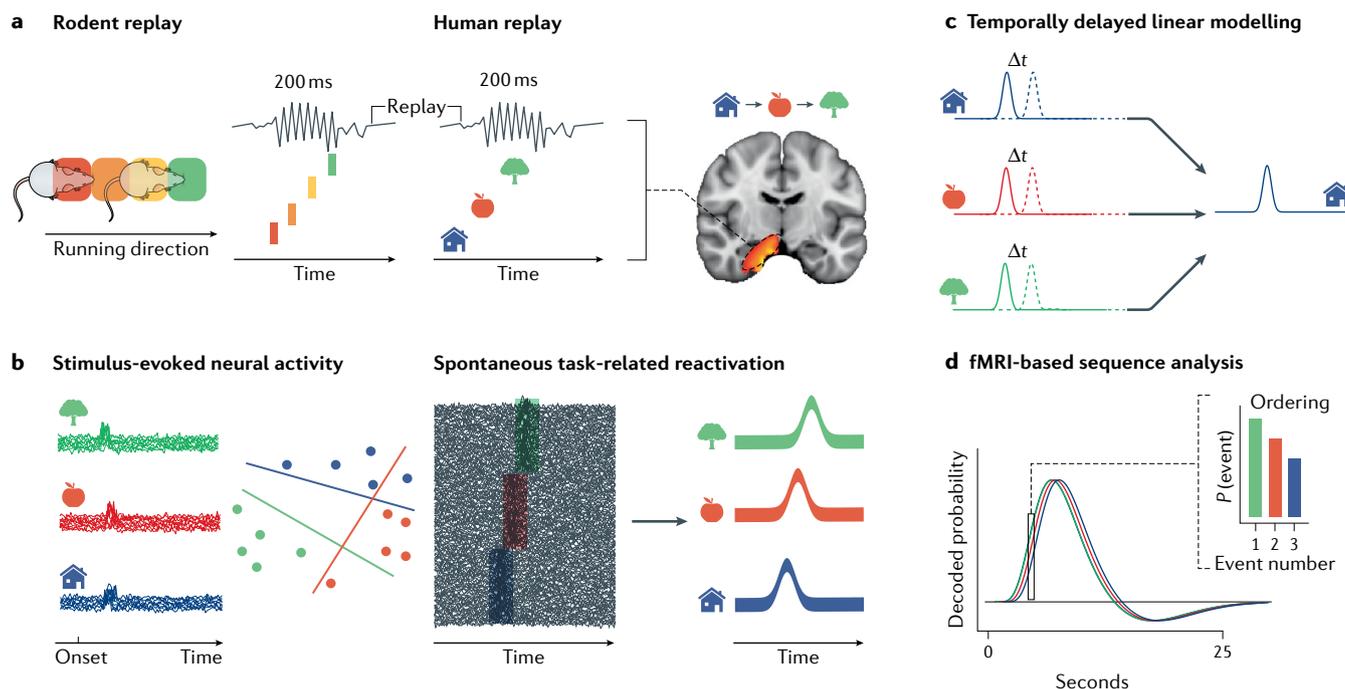


Fig. 2 | ‘Representation-rich’ paradigm of spontaneous neural activity.

**a** | Neural replay in rodents and humans. Rodent replay (left). Different colours indicate firing fields of place cells in the rodent hippocampus. During rest, those place cells are reactivated in a consistent order, termed replay. Rodent replays are time compressed, and are typically associated with sharp wave ripples — the fast oscillations depicted in the figure. Human replay (right). Human replays (of visual stimuli) are also time compressed and are associated with fast oscillations. The brain activation figure shows an initialization of replay events that arises in the hippocampus (pseudocoloured region, where red indicates lower activation and yellow indicates higher activation). **b** | Decoding-based approach, comprising two stages. Stage 1 — indexing neural representations of different task objects. This can be achieved by training a decoding model and finding a multivariate decision boundary in the data pertaining to each object. Examples of three task objects — tree (green), apple (red) and house (blue) — are shown here, with dots indicating samples and lines denoting decision

boundaries (left). Stage 2 — applying these trained decoding models to the spontaneous neural activity of interest enables us to ascertain what has been reactivated and when. Transient bumps in the spontaneous neural activity can then be identified as task-related reactivations for tree (green), apple (red) and house (blue) (right). **c** | Temporally delayed linear modelling approach. It asks whether the averaged statistical likelihood of some transitions (for example, ‘apple’→‘house’) happening is greater than that of others (for example, ‘tree’→‘house’). Dashed lines indicate the time-shifted (by  $\Delta t$ ) copies. **d** | A functional MRI (fMRI)-based sequence analysis approach. Fast sequences of events will cause systematic patterns of overlap in delayed responses that can be inferred<sup>45</sup>. Differently coloured lines indicate the time course of different decoded events. The inset shows an ordering of their reactivation strength for a given time slice. Part **a** is adapted from REF.<sup>23</sup>, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>). Part **c** is adapted from REF.<sup>53</sup>, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>).

that reactivations in both time periods are subject to task-related modulation<sup>62</sup>. While ‘on-task’ reactivations (for example, those occur in inter-trial intervals) relate more directly to immediate task demands, we include these phenomena under the rubric of ‘spontaneous reactivations’ as the representational content in question relates to states that are not immediately determined by current sensory input. For example, mental simulations of future experiences or the recall of episodic memories may be detected ‘on-task’ (that is, during a inter-trial interval or after receipt of an outcome), yet ‘go beyond’ immediate task-evoked processing. Both categories of spontaneous activity can be studied in a similar manner under a ‘representation-rich’ paradigm. We note that studies of human reactivations (both off-task and on-task) are sometimes referred to as ‘replay’ in the literature<sup>58,63</sup>. For consistency, in this Review, we use ‘reactivation’ to refer to task representations obtained during learning and encoding time that are later reinstated in spontaneous neural activity<sup>64</sup>. We use the term ‘replay’ to describe a *sequential* reactivation of these task representations.

**Off-task spontaneous reactivation.** Outside a ‘representation-rich’ approach, studies probing the functional relevance of resting-state activity typically link physiological features (for example, functional connectivity) to behavioural measures of task performance collected before, or after, a resting session. For example, using fMRI, Tambini et al.<sup>65</sup> reported enhanced functional connectivity between the hippocampus and the lateral occipital cortex during rest following an associative memory encoding task, which was related to later memory performance. This approach is analogous to linking electrophysiological signatures of reactivation (for example, SWRs) to memory consolidation in rodents<sup>42</sup>.

Implementing a representation-rich paradigm, other studies have endeavoured to probe cognitive process within resting-state activity<sup>14</sup>. For example, comparison of stimulus-encoding patterns in the hippocampus for the pre-encoding rest period versus the postencoding rest period has revealed a stronger memory reactivation during post-encoding rest that relates to enhanced memory performance<sup>18</sup>. In a decoding analysis of EEG

#### Multivariate decision boundary

A region of a problem space where the output label of a classifier is ambiguous.

**Charles Bonnet syndrome**  
A condition where visual hallucinations occur as a result of vision loss.

signals, previously learnt information was found to be reactivated during sleep at the category level<sup>66</sup>. With use of fMRI data acquired during sleep, reactivation of past memory was found at both the category level<sup>67</sup> and at the level of individual stimuli<sup>68</sup>. More recently, Schapiro et al.<sup>63</sup> showed that prioritized reactivation of weakly encoded memories in the hippocampus during awake rest benefits later memory performance. Together, these studies provide evidence that offline hippocampal reactivation plays an important role in human memory consolidation<sup>69</sup>. This research trend towards use of a representation-rich approach enables fine-grained tracking of the representational content and dynamics of reactivations, mirroring the tracking of spontaneous place cell activity in the rodent hippocampus during rest.

Sequential reactivation (or replay) in humans concerns the ordering of reactivated task states at the representation level. The sequential quality of such reactivation renders it a suitable testbed to assess the neural representation of cognitive map, where only a limited number of transitions are valid (for example, A→B and B→C are valid sequences, but C→A is not, in a linear structure of A→B→C). Schuck and Niv<sup>25</sup> took fMRI-based decoding a step further to provide evidence for sequential reactivation of task states in the human hippocampus during rest. Here, a focus on representational content enabled them to make claims about the regularity and consistency of the same experiences between rest and task, enabling them to connect

hippocampal replay to the neural representation of a mental model in the orbitofrontal cortex<sup>70</sup>. Specifically, they suggested that human replay might participate in building or maintaining a mental representation of the task structure during rest. Relatedly, using MEG at millisecond temporal resolution, Liu et al.<sup>23</sup> demonstrated that organized experiences are sequentially replayed during rest on a fast timescale (40-ms state-to-state transition). Such replays were not mere ‘echoes’ of past experience<sup>35</sup>, but instead were ordered in a manner consistent with a learnt task structure, again suggestive of replay building or maintaining a cognitive map off-task.

Of particular note is that the spontaneous sequences of cortical events, detected in a non-spatial context in humans using MEG<sup>23</sup>, show striking parallels to the characteristics of hippocampal replay in rodents during SWR epochs in spatial tasks<sup>27,31,36</sup>. Like rodent replays, human replays (1) appear spontaneously during rest, (2) compress time from seconds to tens of milliseconds<sup>28</sup>, (3) reverse in direction following receipt of a reward<sup>37</sup>, (4) involve a coordination between the hippocampus and the sensory cortex<sup>32</sup> and (5) are associated with a power increase in ripple frequency (120–150 Hz), which can be source localized to the hippocampus.

**On-task spontaneous reactivation.** A rich literature has characterized the representational content of spontaneous neural activity on-task, ranging from perception<sup>12</sup> to action<sup>71</sup>. An interesting finding is that reactivation of objects in mind can bias subsequent perception, and in an extreme case leads to hallucination. For example, Pajani et al.<sup>12</sup> showed that the representational content of prestimulus activity in the early visual cortex is linked to subsequent perception: if there is a bias towards the expected grating stimuli, it could predispose to perceptual hallucination. Similarly, Hahamy et al.<sup>72</sup> found spontaneous fluctuations in the early visual cortex might activate the visual hierarchy, and drive hallucination in participants with Charles Bonnet syndrome.

Over the past 5 years, there has been an upsurge in representation-rich research on memory and decision-making in the context of RL<sup>73</sup>, especially model-based RL<sup>74</sup>. In RL terms, a ‘model’ details the relationships between current and future states. This has a similar meaning to the notion of ‘cognitive map’. A commonality among these is that a model allows us to infer things we have not experienced directly or explicitly<sup>75</sup>, enabling, for instance, multistep planning<sup>22</sup> or inferential learning<sup>34</sup>. If ‘off-task’ reactivation relates to building or maintaining a mental model of the world, then ‘on-task’ reactivation might be a means to study how such a model is used for adaptive behaviour. The study of model-based reasoning de facto concerns probing internal processes that are not tied to a current sensation, a line of investigation that fits well with ‘representation-rich’ approach.

One important field here is that of memory-based research<sup>73</sup>, which typically focuses on the cued retrieval of associative memories<sup>76</sup>. Studies of associative memory entail an encoding phase (for example, a house → tree association, where the house is the cue, and the tree is the associated event) followed by a cued retrieval phase (for example, house → ?). A pairwise association can be

## Box 2 | RL in human neuroscience

Reinforcement learning (RL) is concerned with a specific family of questions: how to make decisions to maximize the expected future (discounted) cumulative reward (or avoidance of punishment) and how to update or adjust behaviour on the basis of a discrepancy between expectation and experienced outcome (that is, prediction error)<sup>46</sup>. In neuroscience, RL is widely linked to specific neural mechanisms, particularly phasic dopamine signalling in mesolimbic circuits reflecting reward prediction errors<sup>155</sup>.

On the basis of whether RL relies on a mental representation of the task space (that is, relational structures among task states), RL is conventionally divided into model-free and model-based processes<sup>74,156,157</sup>. Model-free RL proceeds via trial-and-error learning and relies on consolidating stimulus–response mappings. The best known model is the Rescorla–Wagner model<sup>158</sup>, which was developed in the context of classical conditioning<sup>159</sup>. Although the Rescorla–Wagner model explains many psychological phenomena, and continues to provide remarkable insights into human learning and decision-making<sup>160</sup>, it does not readily address more flexible forms of cognition, such as those concerned with sequential decision-making or computations that go beyond directly experienced stimulus–response associations (for example, planning detours or considering counterfactuals), both of which necessitate reliance on an internal ‘model’ of the task (that is, model based. Note, model-free method is possible to solve these tasks but is much more inefficient and inflexible).

The ‘two-step’ task is a classical paradigm developed to study model-based RL<sup>156,161–163</sup>. In its original formulation by Daw et al.<sup>162</sup>, this involved two-stage sequential decision-making steps, where only the second-stage choice results in a monetary outcome. The state transition structure between the first stage and the second stage is designed to yield different patterns of choice behaviour in a model-free versus model-based agent (where the former has no internal ‘model’ of the transition structure and makes choices based on cached stimulus–action values). Such tasks may be used to characterize the extent of model-based computation in decision-making, according to the degree to which they make use of an internal model of the task. When combined with the ‘representation-rich’ approach, tasks of this nature yield new insights into the intrinsic neural mechanisms supporting model-based cognition<sup>24,55</sup>. Note that in these neuroimaging studies, the transition structure can also be deterministic<sup>55</sup>, or can even be simplified to a one-step decision<sup>24</sup>, to ease the use of decoding techniques.

**Resting states**

The states when an explicit task is not being performed.

conceived as a simple relational structure, with a cued retrieval framed as memory-based decision-making<sup>73</sup>. For example, using intracranial EEG, Norman et al.<sup>77</sup> found successful memory recall was preceded by an increased probability of hippocampal SWRs, during which there was also a transient re-emergence of activation patterns in higher visual cortical areas. Relatedly, Vaz et al.<sup>76</sup> found bursts of spikes in the human temporal lobe that were organized into sequences during memory encoding, and these same sequences were ‘replayed’ during successful memory retrieval. Using fMRI, Wimmer and Shohamy<sup>60</sup> found monetary rewards led to automatic reactivation of a past associative memory within the hippocampus, with these reactivations biasing later value-based decision-making. This effect was subsequently replicated in a MEG-based version of the same task, using a decoding based technique<sup>57</sup>. More recently, Wimmer et al.<sup>59</sup> studied memory retrieval of more extended episodes consisting of multiple elements, and found a differentiation among retrieval patterns involving clustered representations compared with sequential reactivation of individual episode elements, with more strongly encoded memories retrieved via a clustered reactivation and weaker memories retrieved by sequential reactivation<sup>59</sup>.

Other ‘on-task’ research has focused on activity during decision time or following outcome feedback in value-based decision-making tasks<sup>21,22,55,56,58,61</sup>. In rodents, this is typically studied in a spatial navigation setting<sup>37,78</sup>. In humans, we are able to probe more flexible cognitive processes in an abstract task space using the approach outlined. Here a common experimental design involves model-based RL tasks where participants update the value of each action (in each state) on the basis of experienced or inferred (model-based) rewards, and then make a choice based on these values. For example, when a two-step task was performed in fMRI (BOX 2), prospective reactivation of the task goal was found to support model-based choice<sup>55</sup>, on-task reactivation of counterfactual value signals reflected the non-chosen option<sup>79</sup> and spontaneous reactivation following reward receipt was modulated by prediction error<sup>58</sup>. In the model-based inference, an important role of hippocampal–orbitofrontal interactions is highlighted<sup>80</sup>. Other decision-making studies report spontaneous reorganization of task strategies reflected in neural representations, even before a strategy change is evident in behaviour<sup>81</sup>.

Using a sequentially structured RL task in combination with MEG, Liu et al.<sup>24</sup> identified reverse sequential reactivation (reverse replay) of non-local (that is, inferred) experiences following reward receipt, with a 160-ms state-to-state lag, akin to a putative neural mechanism for model-based RL. This replay was more pronounced for experiences of greater utility for future behaviour, consistent with RL theory<sup>48</sup>. Although this study focused on outcome time, other studies have probed the mental planning process at choice time. In a non-spatial sequential planning task, Kurth-Nelson et al.<sup>22</sup> found fast reverse sequences during planning, these sequences did not reflect the path to be taken, but instead represented equally all valid transitions.

This is suggestive of a process involving mental rehearsal of sequences, possibly to maintain a representation of task structure. Subsequently, Eldar et al.<sup>56</sup> varied both problem complexity and temporal constraints in a decision task, showing that people differ substantially in terms of their decision strategies. Using a different task, Eldar et al.<sup>21</sup> further demonstrated that on-task (during planning) and off-task (during rest) replay supported planning in distinctive ways, dependent on actual decision strategies. Such model-based planning has also been found recently in the aversive domain<sup>61</sup>, akin to findings in the rodent literature<sup>40</sup>.

**Bridging cognition and physiology**

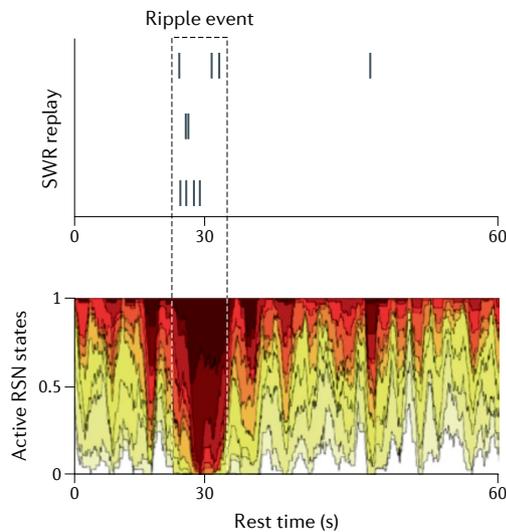
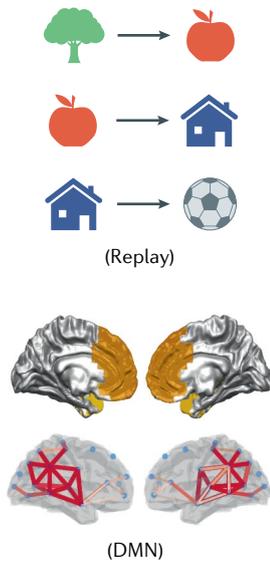
Having considered how a ‘representation-rich’ approach might advance our understanding of human cognition, we now discuss potential insights this approach can bring to studies of resting states, and in particular to psychiatry research, where there has long been a strong emphasis on spontaneous neural activity<sup>2,82,83</sup>.

A predominant focus of resting-state studies is the intrinsic physiological features of brain activity (for example, the covariance of activity between brain regions, termed ‘functional connectivity’). One common approach is to characterize the functional connectivity patterns relating different regions in terms of whole-brain resting-state networks (RSNs)<sup>5,84,85</sup>. Among such RSNs, the DMN<sup>85,86</sup> is of particular interest<sup>87</sup>.

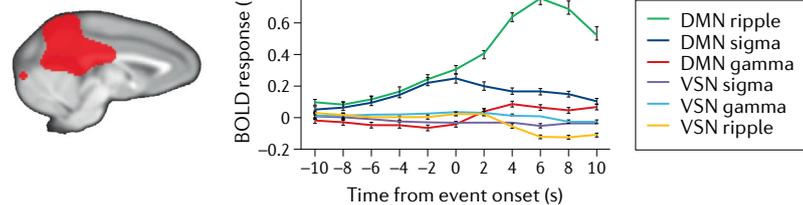
The DMN is a characteristic pattern of interconnected brain regions (which includes the medial prefrontal cortex, posterior cingulate cortex and medial temporal lobes) that show high co-activation off-task<sup>87–89</sup>. Initially, the DMN was thought to be ‘task negative’, as it is typically deactivated during task execution (for example, working memory)<sup>90,91</sup>. Later, brain regions that constitute the DMN were found to be related to internally oriented cognitive states<sup>16</sup>, such as imagination<sup>92</sup>, mind wandering<sup>93</sup>, memory recall<sup>94</sup>, planning<sup>95</sup> or consolidating social information<sup>96</sup>. In recent studies focusing on the neural codes underpinning mental models (for example, conceptual spaces<sup>97</sup>, social spaces<sup>98</sup> or narrative schemas<sup>99</sup>), the distribution of such neural profiles bears a remarkable overlap with DMN, leading to a suggestion that this functional network might be encoding the cognitive map of task space<sup>75</sup>. These interpretations, however, derive in large part from an anatomical overlap with brain regions reported in task-based cognitive studies.

Ideally, we want to link functional connectivity patterns (for example, the DMN) to concurrent spontaneous cognition within the same data set. In a representation-rich approach, this can be achieved by studying the relationship between DMN activations and task state-related reactivations. Combining two recent methodological advances in MEG analysis — measurement of sequential replay during rest<sup>23,53</sup> and tracking of DMN activation dynamics with millisecond temporal resolution<sup>100</sup> — Higgins et al.<sup>15</sup> established a connection between spontaneous human replay and DMN activation in the same resting-state session. More specifically, they showed that human replay exhibits a highly organized temporal structure, where replay events did not

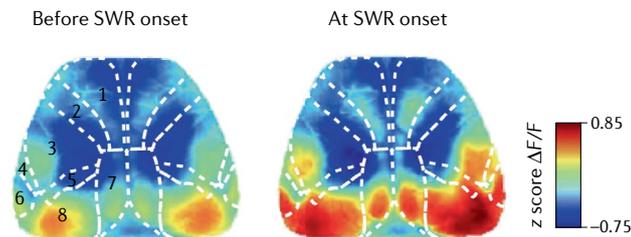
**a Human brain**



**b Monkey brain**



**c Mouse brain**



occur randomly but were instead packaged into transient bursts. The latter coincided with a concentration within epochs of DMN activation (FIG. 3a) characterized by large synchrony in the delta band and/or theta band. Moreover, the DMN was unique among all RSNs in its association with transient increases in higher-frequency power (including the frequency band associated with ripple), which source localized to the temporal lobe. This work suggests that a coupling between temporal-lobe SWR and the DMN might provide a physiological basis for how human replay supports memory consolidation.

A coupling between temporal-lobe SWR and the cortical DMN during rest<sup>15</sup> is reminiscent of prior in vivo work in non-human primates. For example, using simultaneous whole-brain fMRI as well as hippocampal electrophysiology recordings, Kaplan et al.<sup>101</sup> found a selective increase in the DMN following hippocampal ripples, but not other RSNs or hippocampal electrophysiological

**Fig. 3 | Coordination between hippocampal SWRs and large-scale cortical activity across species. a** | In humans, the top part shows an example (60 s in a 5-min resting state) of raster plots (top right) of three replay sequences (top left). The lower part shows the time course of 12 different resting-state network (RSN) reactivations (bottom right, each RSN is indicated by a colour from brown through to yellow). The y axis indicates the reactivation probability of RSNs, where a higher value indicates stronger activation. The blue dots indicate the different regions of interest from where spectral information (for example, the phase) is extracted. The power (yellow) and phase locking (red) profile of the default mode network (DMN) in the human brain, where brighter yellow and darker red indicate higher power and higher phase locking, respectively. The DMN (brown) shows the strongest activation at the time of sharp wave ripples (SWR) replay bursts (bottom left). **b** | In monkeys, hippocampal SWRs have also been shown to be coupled with the DMN (left, red area, measured with functional MRI). This coupling is specific to the DMN (among other RSNs) and ripple frequency (among other frequency bands). **c** | In mice, widespread cortical activation is associated with the onset of hippocampal SWRs. Dashed lines indicate identified cortical regions based on the Allen Brain Atlas (1, secondary motor cortex; 2, primary motor cortex; 3, primary somatosensory cortex; 4, secondary somatosensory cortex; 5, posterior parietal cortex; 6, auditory cortex; 7, retrosplenial cortex; 8, visual cortex). This cortical activity (including the retrosplenial cortex, a part of the DMN) rise occurs immediately before the onset of hippocampal SWRs.  $\Delta F/F$  measures the change in fluorescence intensity relative to its resting level; a higher value indicates higher activation. a.u., arbitrary units; BOLD, blood oxygenation level dependent; VSN, ventral somatomotor network. Part **a** adapted with permission from REF.<sup>15</sup>, Elsevier. Part **b** is adapted from REF.<sup>101</sup>, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>). Part **c** adapted from REF.<sup>102</sup>, Springer Nature Limited.

events (FIG. 3b). Similarly, using simultaneous recordings of hippocampal electrophysiology during wide-field calcium imaging of cortical activity in the mouse brain, Liu et al.<sup>102</sup> found cortical–hippocampal coordination involving hippocampal SWR and medial parietal cortex (part of the DMN) activation (FIG. 3c). This set of results suggests a plausible cross-species function for the DMN during rest in supporting off-task memory consolidation (or map building), potentially through replay in coordination with SWRs.

Contrary to findings in relation to off-task replay, studies of human on-task replay have not shown an association with high-frequency power increases<sup>59</sup>. One intriguing hypothesis is that on-task replay in humans (for example, the slower replay with 160-ms state-to-state time lag)<sup>24</sup> might relate to the theta sequence seen in rodents<sup>103–105</sup> (BOX 1) and might reflect a more ‘conscious’ on-task computation, a rich topic for future work.

**Implications for psychiatry research**

Finally, there are compelling reasons to believe a representation-rich approach can yield novel insights into the neurobiology of psychiatric disorders. Over the last two decades, extensive fMRI, MEG and EEG investigations have reported widespread alterations

**Brownian diffusive spatial trajectories**

Trajectories whose movement is random in space.

**Superdiffusive dynamics**

Random movement but with sudden jumps.

in RSN characteristics in patients with neuropsychiatric disorders compared with healthy volunteers<sup>82,83</sup>. However, in these studies, linkages to clinical variables and cognitive functions are generally inferred indirectly — for example, by relating brain activations to clinical questionnaires<sup>106</sup>. In view of the relationship between spontaneous neural activity and model-based cognition, we consider the ‘representation-rich’ perspective is uniquely placed for bridging a conceptual gap between brain activity, cognition and clinical symptoms in psychiatric disorders.

With this in mind, a condition of particular interest is schizophrenia, a neuropsychiatric disease characterized by symptoms such as delusions, hallucinations and ‘thought disorder’<sup>107</sup>. Genetic mouse models of schizophrenia have identified signatures of abnormal hippocampal reactivations during rest (for example, augmented SWR power and temporal disorganization of place cell reactivations), suggesting that abnormal replay plays a central role in this debilitating condition<sup>108–110</sup>. Patients with a diagnosis of schizophrenia (PScz) exhibit impairments in inferring indirect associations (for example, if  $A > B$  and  $B > C$ , what is the relationship between  $A$  and  $C$ )<sup>111,112</sup>, which can be interpreted as a deficit in leveraging cognitive map. Recently, using an MEG decoding approach, Nour et al.<sup>113</sup> examined neural replay in PScz (versus carefully matched control participants) during rest following a similar relational inference task as in Liu et al.<sup>23</sup>. They found that spontaneous neural replays of learnt task structure, as evident in control participants akin to the findings of Liu et al.<sup>23</sup>, were reduced in PScz. Intriguingly, PScz were characterized by augmented SWR power during replay, and a distorted neural representation of cognitive map, consistent with the genetic mouse model<sup>108</sup>. The sequential replay deficit in PScz related to behavioural impairment in inferring correct sequential relationships between task states, a process likely to require abstracted representation of task structure. These results raise a tentative suggestion that previous reports of DMN deficits in schizophrenia<sup>114</sup> might relate to a compromised maintenance of a correct cognitive map during rest.

This early study of spontaneous neural replay in a clinical population provides motivation for future studies across a range of psychiatric disorders — for example, testing computational hypotheses pertaining to sequential planning deficits (for example, maladaptive pruning) and recurrent intrusive thoughts (for example, obsessions and ruminations) in disorders such as anxiety and depression<sup>115–117</sup>.

**Future directions**

A central goal of neuroscience is to understand how neural activity supports cognition and thereby adaptive behaviour<sup>74,75</sup>. We suggest a relative neglect of cognition in the context of spontaneous neural activity (for example, resting state) can now be redressed by reference to a ‘representation-rich’ approach, and where emerging data indicate it can also inform the study of model-based cognition. We envisage three broad directions where a representation-rich paradigm will advance understanding of cognition in humans and animals alike.

**Sleep and cognition.** An exciting direction now open to investigation is decoding the representational content of sleep. This has broad relevance for understanding both the functional relevance of sleep and its translational implications<sup>118</sup>, given its deficits are shown in most psychiatric disorders.

It has long been hypothesized that memory consolidation, or new memory formation, happens during SWR events in slow wave sleep<sup>119–121</sup>. Early studies of rodent sleep replay showed that a time compression feature<sup>27,122</sup> of neural reactivation supports Hebbian learning by reactivating memory traces within a time window that is amenable to spike timing-dependent plasticity<sup>36,123</sup>. Most recently, evidence of theta sequence expression during rapid eye movement (REM) sleep has been reported in rats after they completed a spatial learning task, suggesting a role for REM-associated theta sequences in memory function as well<sup>124</sup>. In addition, there are interesting differences in replay dynamics during awake rest versus sleep<sup>125,126</sup>. For example, replays in rodents have been found to represent Brownian diffusive spatial trajectories during sleep<sup>127</sup>, while resembling more superdiffusive dynamics during awake rest<sup>126</sup>, hypothesized to serve different computational goals<sup>125</sup>.

In humans, although sleep is a topic of intense investigation, sleep replay has rarely been studied (however, see<sup>66–68</sup> for studies of memory reactivation during sleep). A significant barrier here is the considerable difference in neural signals between sleep and awake<sup>128</sup> states, such that decoding models trained during an awake state may generalize poorly to sleep time<sup>129</sup>. A promising direction is to explore a family of generative models. Thus, instead of treating the mapping between a task variable and multivariate neural activity as a black box (as is the case in discriminative models), this class of model specifies the generative process of neural activity in relation to task variables, a priori. This class of model can, in principle, generalize better from the awake state to sleep if the underlying assumptions are reasonably met<sup>49,130</sup>. These approaches may also allow an exploration of links between sleep replay and dreaming in humans.

**Integrating human and animal neuroscience.** A unique advantage of the ‘representation-rich’ paradigm is its ability to integrate findings across species. This is because neural signals (either from electrophysiology in rodents or from neuroimaging in humans) can be transformed into task-related representations. This focus on the representational level renders possible a comparison of human and animal neuroscience findings.

Previous work on the spatial organization of neural codes for visual objects in the brain, for example, found the inferior temporal cortex supports a common neural representation profile for animate versus inanimate objects across humans and monkeys<sup>131</sup>. Recent work on the temporal profile of spontaneous neural reactivations (for example, replay) suggests that human replay bears a strong resemblance to that seen in rodents<sup>23</sup>. This leads to an expectation of greater crosstalk and assimilation of findings across species under a representation-rich paradigm<sup>132</sup>, especially with use of domain-general methods (such as temporally delayed linear modelling)<sup>53</sup>.

In a recent example of this integrative approach, Barron et al.<sup>133</sup> probed inferential decision-making in both humans (with fMRI) and mice (with electrophysiology) with a similar associative inference task, and revealed hippocampal involvement in both species. In this study, hippocampal replay in rodents represented inferred relationships during rest, whereas the human hippocampus was found to use a prospective code to forecast learnt associations. It is also possible that when solving the same task, distinct species use different cognitive maps, with implications for performance efficiency<sup>134,135</sup>.

**More complex forms of cognition.** Humans possess a remarkable mental ability that extends well beyond spatial cognition, including an ability to reason, to flexibly deploy language and to generalize experience to novel contexts<sup>136,137</sup>. Understanding the neural code of these highly flexible forms of human cognition is of great interest in many related fields, including both neuroscience and artificial intelligence<sup>75</sup>. Under a ‘representation-rich’ approach, it is now possible to probe the internal computations of those complex cognitive process. For example, using MEG-based decoding, Al Roumi et al.<sup>138</sup> studied how sequences, and operations on sequences, are represented in the brain, finding evidence for an abstract, language-like code for flexible sequence representation. Similarly, Liu et al.<sup>23</sup> demonstrated a ‘factorized representation’ (with independent representation of abstract structural knowledge and concrete sensory information) in human replay, which is likely to be useful for inference and generalization in novel contexts<sup>75,139,140</sup>. Undoubtedly there are other organization principles of neural information for supporting flexible behaviour<sup>75,141</sup>. For example,

the dynamics of semantic representations during a visual understanding task might reflect a unique human ability to reason about arbitrary novel problems<sup>142</sup>.

Finally, in the study of abstract psychological processes, it may turn out that there is no apparent decodable content. While decoding mental states, such as emotional states, is possible in principle<sup>143,144</sup>, other, more tractable, approaches, including experience sampling<sup>145,146</sup>, have been used. To gain a complete understanding of the rich dynamics of spontaneous neural activity, it is important that we develop new methods that are suited to probe these abstract cognitive processes in the future.

## Conclusion

Recent advances in decoding cognition from spontaneous neural activity provides a basis for grounding human cognitive studies that are beyond immediate task demand. We suggest a ‘representation-rich’ approach, which relies on the cognitive map of task space, can advance our understanding of a wide range of cognitive processes extending beyond task-evoked response. These processes include memory retrieval, planning and inference, which lie at the heart of sophisticated model-based reasoning. In considering both off-task and on-task neural reactivation, we have outlined how this approach can help bridge a divide between studies of resting states and those that focus on task-evoked activity. Finally, we consider that linking physiological features of neural activity to its representational content will have profound implications for future research in psychiatry, particularly in light of recent findings in schizophrenia.

Published online 8 March 2022

- Uddin, L. Q. Bring the noise: reconceptualizing spontaneous neural activity. *Trends Cogn. Sci.* **24**, 734–746 (2020).
- Zhang, D. & Raichle, M. E. Disease and the brain's dark energy. *Nat. Rev. Neurol.* **6**, 15–28 (2010).
- Becker, R., Van De Ville, D. & Kleinschmidt, A. Alpha oscillations reduce temporal long-range dependence in spontaneous human brain activity. *J. Neurosci.* **38**, 755–764 (2018).
- Allaman, L., Mottaz, A., Kleinschmidt, A. & Guggisberg, A. G. Spontaneous network coupling enables efficient task performance without local task-induced activations. *J. Neurosci.* **40**, 9663–9675 (2020).
- Smith, S. M. et al. Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* **17**, 666–682 (2013).
- Smith, S. M. et al. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl Acad. Sci. USA* **106**, 13040–13045 (2009).
- Tavor, I. et al. Task-free MRI predicts individual differences in brain activity during task performance. *Science* **352**, 216–220 (2016).  
**This study shows that individual differences in brain activity during task performance can be predicted on the basis of its neural profile off-task (that is, during rest).**
- Rudoy, J. D., Voss, J. L., Westerberg, C. E. & Paller, K. A. Strengthening individual memories by reactivating them during sleep. *Science* **326**, 1079–1079 (2009).
- Rasch, B., Büchel, C., Gais, S. & Born, J. Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science* **315**, 1426–1429 (2007).
- Wang, B. et al. Targeted memory reactivation during sleep elicits neural signals related to learning content. *J. Neurosci.* **39**, 6728–6736 (2019).
- Cairney, S. A., El Marj, N. & Staresina, B. P. Memory consolidation is linked to spindle-mediated information processing during sleep. *Curr. Biol.* **28**, 948–954.e4 (2018).
- Pajani, A., Kok, P., Kouider, S. & de Lange, F. P. Spontaneous activity patterns in primary visual cortex predispose to visual hallucinations. *J. Neurosci.* **35**, 12947–12953 (2015).
- Chew, B. et al. Endogenous fluctuations in the dopaminergic midbrain drive behavioral choice variability. *Proc. Natl Acad. Sci. USA* **116**, 18732–18737 (2019).
- Tambini, A. & Davachi, L. Awake reactivation of prior experiences consolidates memories and biases cognition. *Trends Cogn. Sci.* **23**, 876–890 (2019).
- Higgins, C. et al. Replay bursts in humans coincide with activation of the default mode and parietal alpha networks. *Neuron* **109**, 882–893.e7 (2021).  
**This study shows that human replays happen in bursts and are coupled with activation of the DMN, as well as high-frequency power increase in the temporal lobe.**
- Yeshurun, Y., Nguyen, M. & Hasson, U. The default mode network: where the idiosyncratic self meets the shared social world. *Nat. Rev. Neurosci.* **22**, 181–192 (2021).
- Sutherland, G. R. & McNaughton, B. Memory trace reactivation in hippocampal and neocortical neuronal ensembles. *Curr. Opin. Neurobiol.* **10**, 180–186 (2000).
- Tambini, A. & Davachi, L. Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proc. Natl Acad. Sci. USA* **110**, 19591–19596 (2013).
- Kriegeskorte, N., Mur, M. & Bandettini, P. A. Representational similarity analysis—connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
- Diedrichsen, J. & Kriegeskorte, N. Representational models: a common framework for understanding encoding, pattern-component, and representational-similarity analysis. *PLoS Comput. Biol.* **13**, e1005508 (2017).
- Eldar, E., Lièvre, G., Dayan, P. & Dolan, R. J. The roles of online and offline replay in planning. *eLife* **9**, e56911 (2020).
- Kurth-Nelson, Z., Economides, M., Dolan, R. J. & Dayan, P. Fast sequences of non-spatial state representations in humans. *Neuron* **91**, 194–204 (2016).
- Liu, Y., Dolan, R. J., Kurth-Nelson, Z. & Behrens, T. E. J. Human replay spontaneously reorganizes experience. *Cell* **178**, 640–652 (2019).  
**This study shows that human replay (measured with MEG) exhibits similarities to rodent replay, while representing sensory and structural information independently, facilitating generalization in a novel context.**
- Liu, Y., Mattar, M. G., Behrens, T. E., Daw, N. D. & Dolan, R. J. Experience replay is associated with efficient nonlocal learning. *Science* **372**, eabf1357 (2021).
- Schuck, N. W. & Niv, Y. Sequential replay of nonspatial task states in the human hippocampus. *Science* **364**, eaaw5181 (2019).  
**This study shows that human replay (measured with fMRI) can be captured in the hippocampus and is related to maintaining a neural representation of task space in the orbitofrontal cortex.**
- Burgess, N. & O'Keefe, J. Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus* **6**, 749–762 (1996).
- Skaggs, W. E. & McNaughton, B. L. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science* **271**, 1870–1873 (1996).
- Nádasy, Z., Hirase, H., Czurkó, A., Csicsvari, J. & Buzsáki, G. Replay and time compression of recurring

- spike sequences in the hippocampus. *J. Neurosci.* **19**, 9497–9507 (1999).
29. Louie, K. & Wilson, M. A. Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron* **29**, 145–156 (2001).
  30. Foster, D. J. & Wilson, M. A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440**, 680 (2006).
  31. Diba, K. & Buzsáki, G. Forward and reverse hippocampal place-cell sequences during ripples. *Nat. Neurosci.* **10**, 1241 (2007).
  32. Ji, D. & Wilson, M. A. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat. Neurosci.* **10**, 100 (2007).
  33. Davidson, T. J., Kloosterman, F. & Wilson, M. A. Hippocampal replay of extended experience. *Neuron* **63**, 497–507 (2009).
  34. Karlsson, M. P. & Frank, L. M. Awake replay of remote experiences in the hippocampus. *Nat. Neurosci.* **12**, 913 (2009).
  35. Gupta, A. S., van der Meer, M. A., Touretzky, D. S. & Redish, A. D. Hippocampal replay is not a simple function of experience. *Neuron* **65**, 695–705 (2010).
  36. Carr, M. F., Jadhav, S. P. & Frank, L. M. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat. Neurosci.* **14**, 147 (2011).
  37. Ambrose, R. E., Pfeiffer, B. E. & Foster, D. J. Reverse replay of hippocampal place cells is uniquely modulated by changing reward. *Neuron* **91**, 1124–1136 (2016).
  38. Ólafsdóttir, H. F., Carpenter, F. & Barry, C. Coordinated grid and place cell replay during rest. *Nat. Neurosci.* **19**, 792 (2016).
  39. O'Neill, J., Boccarda, C. N., Stella, F., Schoenenberger, P. & Csicsvari, J. Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. *Science* **355**, 184–188 (2017).
  40. Wu, C.-T., Haggerty, D., Kemere, C. & Ji, D. Hippocampal awake replay in fear memory retrieval. *Nat. Neurosci.* **20**, 571–580 (2017).
  41. Buzsáki, G. Hippocampal sharp wave-ripple: a cognitive biomarker for episodic memory and planning. *Hippocampus* **25**, 1073–1188 (2015). **This study shows that hippocampal SWRs are implicated in many cognitive processes, the deficit of which occurs in certain psychiatric conditions, such as schizophrenia.**
  42. Jadhav, S. P., Kemere, C., German, P. W. & Frank, L. M. Awake hippocampal sharp-wave ripples support spatial memory. *Science* **336**, 1454–1458 (2012).
  43. Foster, D. J. Replay comes of age. *Annu. Rev. Neurosci.* **40**, 581–602 (2017).
  44. Ólafsdóttir, H. F., Bush, D. & Barry, C. The role of hippocampal replay in memory and planning. *Curr. Biol.* **28**, R37–R50 (2018).
  45. Wittkuhn, L. & Schuck, N. W. Dynamics of fMRI patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nat. Commun.* **12**, 1795 (2021).
  46. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 2018).
  47. Sutton, R. S. Dyna, an integrated architecture for learning, planning, and reacting. *ACM Sigart Bull.* **2**, 160–163 (1991).
  48. Mattar, M. G. & Daw, N. D. Prioritized memory access explains planning and hippocampal replay. *Nat. Neurosci.* **21**, 1609 (2018).
  49. Guggenmos, M., Sterzer, P. & Cichy, R. M. Multivariate pattern analysis for MEG: a comparison of dissimilarity measures. *NeuroImage* **173**, 434–447 (2018).
  50. Norman, K. A., Polyn, S. M., Detre, G. J. & Haxby, J. V. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* **10**, 424–430 (2006).
  51. Haxby, J. V. et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425–2430 (2001).
  52. Peelen, M. V. & Downing, P. E. Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations. *Trends Cogn. Sci.* **11**, 4–5 (2007).
  53. Liu, Y. et al. Temporally delayed linear modelling (TDLM) measures replay in both animals and humans. *eLife* **10**, e66917 (2021). **This study shows that fast neural sequences of spontaneous reactivations can be captured non-invasively in the human brain, and the same method can be used to quantify rodent replay in electrophysiology recordings.**
  54. Harris, K. D. Nonsense correlations in neuroscience. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.11.29.402719> (2020).
  55. 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 (2015).
  56. Eldar, E., Bae, G. J., Kurth-Nelson, Z., Dayan, P. & Dolan, R. J. Magnetoencephalography decoding reveals structural differences within integrative decision processes. *Nat. Hum. Behav.* **2**, 670–681 (2018).
  57. Kurth-Nelson, Z., Barnes, G., Sejdic, D., Dolan, R. & Dayan, P. Temporal structure in associative retrieval. *eLife* **4**, e04919 (2015).
  58. Momennejad, I., Otto, A. R., Daw, N. D. & Norman, K. A. Offline replay supports planning in human reinforcement learning. *eLife* **7**, e32548 (2018).
  59. Wimmer, G. E., Liu, Y., Vehar, N., Behrens, T. E. J. & Dolan, R. J. Episodic memory retrieval success is associated with rapid replay of episode content. *Nat. Neurosci.* **23**, 1025–1033 (2020).
  60. Wimmer, G. E. & Shohamy, D. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* **338**, 270–273 (2012).
  61. Wise, T., Liu, Y., Chowdhury, F. & Dolan, R. J. Model-based aversive learning in humans is supported by preferential task state reactivation. *Sci. Adv.* **7**, eabf9616 (2021).
  62. Wittkuhn, L., Chien, S., Hall-McMaster, S. & Schuck, N. W. Replay in minds and machines. *Neurosci. Biobehav. Rev.* **129**, 367–388 (2021).
  63. Schapiro, A. C., McDevitt, E. A., Rogers, T. T., Mednick, S. C. & Norman, K. A. Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nat. Commun.* **9**, 3920 (2018).
  64. Genzel, L. et al. A consensus statement: defining terms for reactivation analysis. *Phil. Trans. R. Soc. B* **375**, 202000001 (2020).
  65. Tambini, A., Keiz, N. & Davachi, L. Enhanced brain correlations during rest are related to memory for recent experiences. *Neuron* **65**, 280–290 (2010).
  66. Schönauer, M. et al. Decoding material-specific memory reprocessing during sleep in humans. *Nat. Commun.* **8**, 15404 (2017).
  67. Shanahan, L. K., Gjorgjieva, E., Paller, K. A., Kahnt, T. & Gottfried, J. A. Odor-evoked category reactivation in human ventromedial prefrontal cortex during sleep promotes memory consolidation. *eLife* **7**, e39681 (2018).
  68. Deuker, L. et al. Memory consolidation by replay of stimulus-specific neural activity. *J. Neurosci.* **33**, 19373–19383 (2013).
  69. Antony, J. W. & Schapiro, A. C. Active and effective replay: systems consolidation reconsidered again. *Nat. Rev. Neurosci.* **20**, 506–507 (2019).
  70. Schuck, N. W. et al. Human orbitofrontal cortex represents a cognitive map of state space. *Neuron* **91**, 1402–1412 (2016).
  71. Kornysheva, K. et al. Neural competitive queuing of ordinal structure underlies skilled sequential action. *Neuron* **101**, 1166–1180.e3 (2019).
  72. Hahamy, A., Wilf, M., Rosin, B., Behrmann, M. & Malach, R. How do the blind 'see'? The role of spontaneous brain activity in self-generated perception. *Brain* **144**, 340–353 (2021).
  73. Shohamy, D. & Daw, N. D. Integrating memories to guide decisions. *Curr. Opin. Behav. Sci.* **5**, 85–90 (2015).
  74. Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80**, 312–325 (2013). **This article provides a road map for the neuroscience research of decision-making, with implications for the 'representation-rich' paradigm.**
  75. Behrens, T. E. J. et al. What is a cognitive map? organizing knowledge for flexible behavior. *Neuron* **100**, 490–509 (2018). **This article discusses the functional relevance of a cognitive map, especially in terms of its computational role in human cognition as well as artificial intelligence.**
  76. Vaz, A. P., Wittig, J. H., Inati, S. K. & Zaghoul, K. A. Replay of cortical spiking sequences during human memory retrieval. *Science* **367**, 1131–1134 (2020).
  77. Norman, Y. et al. Hippocampal sharp-wave ripples linked to visual episodic recollection in humans. *Science* **365**, eaax1030 (2019).
  78. Ólafsdóttir, H. F., Barry, C., Saleem, A. B., Hassabis, D. & Spiers, H. J. Hippocampal place cells construct reward related sequences through unexplored space. *eLife* **4**, e06063 (2015).
  79. Moneta, N., Garvert, M. M., Heekeren, H. R. & Schuck, N. W. Parallel representation of context and multiple context-dependent values in ventro-medial prefrontal cortex. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.05.17.435844> (2021).
  80. Wang, F., Schoenbaum, G. & Kahnt, T. Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PLoS Biol.* **18**, e3000578 (2020).
  81. Schuck, N. W. et al. Medial prefrontal cortex predicts internally driven strategy shifts. *Neuron* **86**, 331–340 (2015).
  82. Woodward, N. D. & Cascio, C. J. Resting-state functional connectivity in psychiatric disorders. *JAMA Psychiatry* **72**, 743–744 (2015).
  83. Snyder, A. Z. & Raichle, M. E. A brief history of the resting state: the Washington University perspective. *Neuroimage* **62**, 902–910 (2012).
  84. Smith, S. M. et al. Resting-state fMRI in the human connectome project. *Neuroimage* **80**, 144–168 (2013).
  85. Raichle, M. E. & Snyder, A. Z. A default mode of brain function: a brief history of an evolving idea. *Neuroimage* **37**, 1083–1090 (2007).
  86. Gusnard, D. A. & Raichle, M. E. Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* **2**, 685–694 (2001).
  87. Raichle, M. E. The brain's default mode network. *Annu. Rev. Neurosci.* **38**, 433–447 (2015).
  88. Buckner, R. L. & DiNicola, L. M. The brain's default network: updated anatomy, physiology and evolving insights. *Nat. Rev. Neurosci.* **20**, 593–608 (2019).
  89. Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R. & Buckner, R. L. Functional-anatomic fractionation of the brain's default network. *Neuron* **65**, 550–562 (2010).
  90. Raichle, M. E. et al. A default mode of brain function. *Proc. Natl Acad. Sci. USA* **98**, 676–682 (2001).
  91. Buckner, R. L. The serendipitous discovery of the brain's default network. *Neuroimage* **62**, 1137–1145 (2012).
  92. Agnati, L. F., Guidolin, D., Battistin, L., Pagnoni, G. & Fuxe, K. The neurobiology of imagination: possible role of interaction-dominant dynamics and default mode network. *Front. Psychol.* **4**, 296 (2013).
  93. Smallwood, J. & Schooler, J. W. The science of mind wandering: empirically navigating the stream of consciousness. *Annu. Rev. Psychol.* **66**, 487–518 (2015).
  94. Schacter, D. L., Addis, D. R. & Buckner, R. L. Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* **8**, 657–661 (2007).
  95. Hassabis, D. & Maguire, E. A. Deconstructing episodic memory with construction. *Trends Cogn. Sci.* **11**, 299–306 (2007).
  96. Meyer, M. L., Davachi, L., Ochsner, K. N. & Lieberman, M. D. Evidence that default network connectivity during rest consolidates social information. *Cereb. Cortex* **29**, 1910–1920 (2019).
  97. Constantinescu, A. O., O'Reilly, J. X. & Behrens, T. E. J. Organizing conceptual knowledge in humans with a gridlike code. *Science* **352**, 1464 (2016). **This study provides evidence for grid-like coding of non-physical space in the brain areas that constitute the DMN.**
  98. Park, S. A., Miller, D. S. & Boorman, E. D. Novel inferences in a multidimensional social network use a grid-like code. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.05.29.124651> (2020).
  99. Baldassano, C., Hasson, U. & Norman, K. A. Representation of real-world event schemas during narrative perception. *J. Neurosci.* **38**, 9689–9699 (2018).
  100. Vidaurre, D., Myers, N. E., Stokes, M., Nobre, A. C. & Woolrich, M. W. Temporally unconstrained decoding reveals consistent but time-varying stages of stimulus processing. *Cereb. Cortex* **29**, 863–874 (2019).
  101. Kaplan, R. et al. Hippocampal sharp-wave ripples influence selective activation of the default mode network. *Curr. Biol.* **26**, 686–691 (2016). **This study shows a selective coupling between hippocampal ripples and activation of the DMN in the monkey brain.**
  102. Liu, X. et al. Multimodal neural recordings with Neuro-FITM uncover diverse patterns of cortical-hippocampal interactions. *Nat. Neurosci.* **24**, 886–896 (2021).
  103. Wang, M., Foster, D. J. & Pfeiffer, B. E. Alternating sequences of future and past behavior encoded within hippocampal theta oscillations. *Science* **370**, 247 (2020).

104. Buzsáki, G. & Moser, E. I. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.* **16**, 130 (2013).  
**This article provides a synthesis of the hippocampal-entorhinal system in supporting both navigation and memory.**
105. Redish, A. D. Vicarious trial and error. *Nat. Rev. Neurosci.* **17**, 147 (2016).
106. Sylvester, C. M. et al. Individual-specific functional connectivity of the amygdala: a substrate for precision psychiatry. *Proc. Natl Acad. Sci. USA* **117**, 3808–3818 (2020).
107. McCutcheon, R. A., Marques, T. R. & Howes, O. D. Schizophrenia — an overview. *JAMA Psychiatry* **77**, 201–210 (2020).
108. Suh, J., Foster, D. J., Davoudi, H., Wilson, M. A. & Tonegawa, S. Impaired hippocampal ripple-associated replay in a mouse model of schizophrenia. *Neuron* **80**, 484–493 (2013).
109. Altman, C., Harrold, J., Jaaro-Peled, H., Sawa, A. & Foster, D. J. Disordered ripples are a common feature of genetically distinct mouse models relevant to schizophrenia. *Mol. Neuropsychiatry* **1**, 52–59 (2015).
110. Zeidman, P. & Maguire, E. A. Anterior hippocampus: the anatomy of perception, imagination and episodic memory. *Nat. Rev. Neurosci.* **17**, 173–182 (2016).
111. Adams, R. A. et al. Impaired theta phase coupling underlies frontotemporal dysconnectivity in schizophrenia. *Brain* **143**, 1261–1277 (2020).
112. Titone, D., Ditman, T., Holzman, P. S., Eichenbaum, H. & Levy, D. L. Transitive inference in schizophrenia: impairments in relational memory organization. *Schizophr. Res.* **68**, 235–247 (2004).
113. Nour, M. M., Liu, Y., Arumham, A., Kurth-Nelson, Z. & Dolan, R. Impaired neural replay of inferred relationships in schizophrenia. *Cell* **184**, 4315–4328 (2021).  
**This study is the first demonstration of augmented ripple power, cognitive map deficit and diminished replay in PScz.**
114. Whitfield-Gabrieli, S. & Ford, J. M. Default mode network activity and connectivity in psychopathology. *Annu. Rev. Clin. Psychol.* **8**, 49–76 (2012).
115. Huys, Q. J. et al. Interplay of approximate planning strategies. *Proc. Natl Acad. Sci. USA* **112**, 3098–3103 (2015).
116. Huys, Q. J. et al. Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Comput. Biol.* **8**, e1002410 (2012).
117. Heller, A. S. & Bagot, R. C. Is hippocampal replay a mechanism for anxiety and depression? *JAMA Psychiatry* **77**, 431–432 (2020).
118. Lewis, P. A., Knoblich, G. & Poe, G. How memory replay in sleep boosts creative problem-solving. *Trends Cogn. Sci.* **22**, 491–503 (2018).
119. Klinzing, J. C., Niethard, N. & Born, J. Mechanisms of systems memory consolidation during sleep. *Nat. Neurosci.* **22**, 1598–1610 (2019).
120. Wei, Y., Krishnan, G. P., Marshall, L., Martinetz, T. & Bazhenov, M. Stimulation augments spike sequence replay and memory consolidation during slow-wave sleep. *J. Neurosci.* **40**, 811–824 (2020).
121. Tamminen, J., Lambon Ralph, M. A. & Lewis, P. A. The role of sleep spindles and slow-wave activity in integrating new information in semantic memory. *J. Neurosci.* **33**, 15376–15381 (2013).
122. Wilson, M. A. & McNaughton, B. L. Reactivation of hippocampal ensemble memories during sleep. *Science* **265**, 676–679 (1994).
123. Pfeiffer, B. E. The content of hippocampal “replay”. *Hippocampus* **30**, 6–18 (2020).  
**Human replay (measured with MEG) is found to bear many similarities to rodent replay, and here is shown to represent sensory and structural information independently, thus facilitating generalization in a novel context.**
124. Zielinski, M. C., Shin, J. D. & Jadhav, S. P. Hippocampal theta sequences in REM sleep during spatial learning. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.04.15.439854> (2021).
125. McNamee, D. C., Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. Flexible modulation of sequence generation in the entorhinal–hippocampal system. *Nat. Neurosci.* **24**, 851–862 (2021).
126. Krause, E. L. & Drugowitsch, J. A large majority of awake hippocampal sharp-wave ripples feature spatial trajectories with momentum. *Neuron* **110**, 722–733.e8 (2021).
127. Stella, F., Baracska, P., O’Neill, J. & Csicsvari, J. Hippocampal reactivation of random trajectories resembling Brownian diffusion. *Neuron* **102**, 450–461.e7 (2019).
128. Pereira, S. I. R. & Lewis, P. A. Sleeping through brain excitation and inhibition. *Nat. Neurosci.* **23**, 1037–1039 (2020).
129. Belal, S. et al. Identification of memory reactivation during sleep by EEG classification. *Neuroimage* **176**, 203–214 (2018).
130. Higgins, C. *Uncovering Temporal Structure in Neural Data with Statistical Machine Learning Models* (Univ. Oxford, 2019).
131. Kriegeskorte, N. et al. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* **60**, 1126–1141 (2008).
132. Barron, H. C., Mars, R. B., Dupret, D., Lerch, J. P. & Sampaio-Baptista, C. Cross-species neuroscience: closing the explanatory gap. *Phil. Trans. R. Soc. B* **376**, 20190633 (2021).
133. Barron, H. C. et al. Neuronal computation underlying inferential reasoning in humans and mice. *Cell* **183**, 228–243.e21 (2020).
134. Akam, T., Costa, R. & Dayan, P. Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. *PLoS Comput. Biol.* **11**, e1004648 (2015).
135. Miranda, B., Malalasekera, W. M. N., Behrens, T. E., Dayan, P. & Kennerly, S. W. Combined model-free and model-sensitive reinforcement learning in non-human primates. *PLoS Comput. Biol.* **16**, e1007944 (2020).
136. Mack, M. L., Preston, A. R. & Love, B. C. Ventromedial prefrontal cortex compression during concept learning. *Nat. Commun.* **11**, 46 (2020).
137. Morton, N. W., Schlichting, M. L. & Preston, A. R. Representations of common event structure in medial temporal lobe and frontoparietal cortex support efficient inference. *Proc. Natl Acad. Sci. USA* **117**, 29338–29345 (2020).
138. Al Roumi, F., Marti, S., Wang, L., Amalric, M. & Dehaene, S. Mental compression of spatial sequences in human working memory using numerical and geometrical primitives. *Neuron* **109**, 2627–2639.e4 (2021).
139. Bernardi, S. et al. The geometry of abstraction in hippocampus and prefrontal cortex. *Cell* **183**, 954–967.e21 (2020).
140. Higgins, I. et al. Darla: improving zero-shot transfer in reinforcement learning. *Proc. Int. Conf. Mach. Learn.* **70**, 1480–1490 (2017).
141. Whittington, J. C. et al. The Tolman-Eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. *Cell* **183**, 1249–1263.e23 (2020).  
**This study describes a unifying computational model of the hippocampal–entorhinal system for generalization and inference in an arbitrary relational graph (including physical space and memory).**
142. Schwartenbeck, P. et al. Generative replay for compositional visual understanding in the prefrontal-hippocampal circuit. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.06.06.447249> (2021).
143. Kragel, P. A., Knodt, A. R., Hariri, A. R. & LaBar, K. S. Decoding spontaneous emotional states in the human brain. *PLoS Biol.* **14**, e2000106 (2016).
144. Tusche, A., Smallwood, J., Bernhardt, B. C. & Singer, T. Classifying the wandering mind: revealing the affective content of thoughts during task-free rest periods. *Neuroimage* **97**, 107–116 (2014).
145. Van Calster, L., D’Argembeau, A., Salmon, E., Peters, F. & Majerus, S. Fluctuations of attentional networks and default mode network during the resting state reflect variations in cognitive states: evidence from a novel resting-state experience sampling method. *J. Cogn. Neurosci.* **29**, 95–113 (2017).
146. Smallwood, J. et al. The neural correlates of ongoing conscious thought. *iScience* **24**, 102132 (2021).
147. Colgin, L. L. Rhythms of the hippocampal network. *Nat. Rev. Neurosci.* **17**, 239 (2016).
148. Olafsdóttir, H. F., Carpenter, F. & Barry, C. Task demands predict a dynamic switch in the content of awake hippocampal replay. *Neuron* **96**, e926 (2017).
149. Gillespie, A. K. et al. Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron* **109**, 3149–3163.e6 (2021).
150. Chadwick, A., van Rossum, M. C. & Nolan, M. F. Independent theta phase coding accounts for CA1 population sequences and enables flexible remapping. *eLife* **4**, e03542 (2015).
151. Skaggs, W. E., McNaughton, B. L., Wilson, M. A. & Barnes, C. A. Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* **6**, 149–172 (1996).
152. Wikenheiser, A. M. & Redish, A. D. Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* **18**, 289–294 (2015).
153. Johnson, A. & Redish, A. D. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* **27**, 12176–12189 (2007).
154. Kay, K. et al. Constant sub-second cycling between representations of possible futures in the hippocampus. *Cell* **180**, 552–567.e25 (2020).
155. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
156. Doll, B. B., Simon, D. A. & Daw, N. D. The ubiquity of model-based reinforcement learning. *Curr. Opin. Neurobiol.* **22**, 1075–1081 (2012).
157. Daw, N. D. & Dayan, P. The algorithmic anatomy of model-based evaluation. *Phil. Trans. R. Soc. B* **369**, 20130478 (2014).
158. Siegel, S. & Allan, L. G. The widespread influence of the Rescorla-Wagner model. *Psychon. Bull. Rev.* **3**, 314–321 (1996).
159. Sutton, R. S. & Barto, A. G. in *Proceedings of the Ninth Annual Conference of the Cognitive Science Society* (1987).
160. Gallistel, C. R., LoLordo, V. M., Rozin, P. & Seligman, M. E. P. Robert A. Rescorla (1940–2020). *Am. Psychol.* **76**, 391–392 (2021).
161. Vikbladh, O. M. et al. Hippocampal contributions to model-based planning and spatial memory. *Neuron* **102**, 683–693.e4 (2019).
162. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans’ choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
163. Kool, W., Cushman, F. A. & Gershman, S. J. When does model-based control pay off? *PLoS Comput. Biol.* **12**, e1005090 (2016).

**Acknowledgements**

The authors thank E. Wimmer, C. Higgins and P. Schwartenbeck for helpful discussions and fruitful collaborations regarding the work and ideas presented in this Review. This work was supported by the Fundamental Research Funds for the Central Universities (to Y.L.), a Wellcome Trust Investigator Award (098362/Z/12/Z to R.J.D.), a UCL Welcome PhD Fellowship for Clinicians (102186/B/13/Z to M.M.N.), a Wellcome Trust Senior Research Fellowship (104765/Z/14/Z to T.B.), a Principal Research Fellowship (219525/Z/19/Z to T.B.), a James S. McDonnell Foundation Award (JSMF220020372 to T.B.), an Independent Research Group Grant from the Max Planck Society (M.TN.A.BILD0004 to N.W.S.) and a Starting Grant from the European Union (ERC-2019-StG REPLAY-852669 to N.W.S.). M.M.N. is a pre-doctoral fellow of the International Max Planck Research School on Computational Methods in Psychiatry and Ageing Research (<https://www.mps-ucl-centre.mpg.de/en/comp2psych>). The Max Planck UCL Centre for Computational Psychiatry and Ageing Research is supported by UCL and the Max Planck Society. The Wellcome Centre for Human Neuroimaging is supported by core funding from the Wellcome Trust (203147/Z/16/Z). The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z).

**Author contributions**

Y.L. researched data for article and contributed substantially to discussion of the content, writing, review and editing of the manuscript before submission. T.E.J.B. contributed substantially to discussion of the content of the manuscript and contributed to the writing, review and editing of the manuscript. R.J.D. contributed to the writing, review and editing of the manuscript. M.M.N. and N.W.S. contributed to the writing of the manuscript.

**Competing interests**

The authors declare no competing interests.

**Peer review information**

*Nature Reviews Neuroscience* thanks J. Andrews-Hanna and the other anonymous reviewer(s) for their contribution to the peer review of this work.

**Publisher’s note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2022