



Opinion piece



Cite this article: Nalluru S, Glitz L, Przygodda X, Mancini V, Barron HC. 2026 The hippocampus as a generative model. *Phil. Trans. R. Soc. B* **381**: 20250247.

<https://doi.org/10.1098/rstb.2025.0247>

Received: 2 November 2025

Accepted: 25 March 2026

One contribution of 18 to a theme issue 'The role of hippocampal predictions in cognition: bridging perception and memory'.

Subject Areas:

neuroscience, physiology, behaviour, cognition

Keywords:

hippocampus, memory, generative model, latent cause, sharp-wave ripple, theta

Author for correspondence:

Helen C. Barron

e-mail: helen.barron@ndcn.ox.ac.uk

[†]These authors contributed equally to the study.

The hippocampus as a generative model

Sumedha Nalluru^{1,3,2,†}, Leonie Glitz^{1,3,2,4,†}, Xenia Przygodda^{1,3,2},
Valentina Mancini^{1,3,2,5} and Helen C. Barron^{1,3,2}

¹Medical Research Council Centre of Research Excellence in Restorative Neural Dynamics, Nuffield Department for Clinical Neurosciences, and ²Oxford Centre for Integrative Neuroimaging, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK

³Brain Network Dynamics Unit, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX1 3TH, UK

⁴Department of Experimental Psychology, University of Oxford, Oxford OX1 3EL, UK

⁵Department of Psychiatry, University of Oxford, Oxford OX3 7JX, UK

VM, 0000-0003-4411-896X; HCB, 0000-0002-4575-6472

A generative model can be defined as a model of the latent causes of sensory input that can be used to generate new data samples. By examining empirical evidence and computational theory, we propose that the hippocampus can be characterized as a generative model. The hippocampus is a brain region important for memory. Recordings of neural activity from the hippocampus have led to the view that the hippocampus represents a cognitive map by abstracting a low-dimensional representation of the external world. We extend this view to suggest the hippocampus represents the latent, unobserved causes of sensory data by virtue of the position of the hippocampus within the deep cortical hierarchy. These representations of unobserved latent causes endow the hippocampus with capacity to generate new data samples that allow exploration of future hypotheticals and provide an internally generated training signal back to the generative model. We explore how perturbations to the hippocampal generative model may explain core symptoms of neuropsychiatric disorders such as those observed in psychosis. Together, this perspective provides a unified account of hippocampal function that explains how computations performed by the hippocampus support higher-order cognition and adaptive behaviour.

This article is part of the theme issue 'The role of hippocampal predictions in cognition: bridging perception and memory'.

1. What is a generative model?

Generative models have long been used to explain complex behaviours [1–3]. However, the neural circuit implementation is only beginning to be fully realized via an ongoing dialogue between artificial intelligence (AI) and neuroscience. Here, we propose that the hippocampus, a brain region important for memory, may be conceptualized as representing the apex of a generative model that is represented by the cortical hierarchy. Before examining evidence to support this view, we first provide a conceptual and computational definition of a generative model.

To define what we mean by a generative model, first we consider how the brain receives sensory input and uses this input to guide perception and behaviour. For example, a child may see a banana from different locations. While the exact visual input is not memorized, each time the child sees the banana, visual input provides training data to extract and learn a hidden (or 'latent') state (in this case, 'banana'). This latent state is unchanging despite variation in the sensory input, illustrating how perception itself requires inference and is inherently constructive. Beyond perception, more abstract latent states can be inferred. For example, a child may have to leave the house

to infer that it is raining (wet pavement, puddles, etc.), yet over time, a diverse set of sensory inputs (raindrops on the window, pitter-patter sound outside) provide training data to extract and learn the hidden (or 'latent') state (in this case, 'rainy day'). By representing latent states, the brain is endowed with capacity to generate and imagine entirely new examples (e.g. a new banana, rain tomorrow). These examples illustrate the core function of a generative model: past experiences provide training data to learn and later infer hidden (or 'latent') rules that can be used to generate new outputs.

From a computational perspective, a generative model can be defined as a model of the causes and states that generate observed sensory data. These causes and states are *not directly observed* but are instead *inferred*, and therefore described as 'latent', 'hidden' or 'unobserved'. We refer to the process of inferring latent causes from sensory data as *perceptual inference*. Perceptual inference can be mathematically formalized using a *Bayesian framework* [4,5]. Given sensory data X , the probability of latent cause Z is described as the *posterior* and formalized as: $p(Z|X)$. The posterior is estimated by combining the *likelihood* of observing the sensory data ($p(X|Z)$) with *prior* beliefs ($p(Z)$) that are held about the environment.

In addition to perceptual inference, the generative model can be used to *generate new data samples*, by calculating the probability of sensory data X given the model's representation of the latent cause Z (i.e. $p(X|Z)$). Thus, latent state representations are used to reconstruct or simulate sensory experience. Generating new data samples can serve several purposes. First, new data can be used to *predict future observations*. These predictions can be defined as the model's internally generated expectations of sensory data, given the assumptions or beliefs about the world represented by the generative model. By predicting future observations, the generative model may explain incoming sensory data before they arrive, to minimize surprise, or even represent future hypotheticals to guide planning and inform decision-making. When these predictions extend beyond past experience, the generative model can be used to *infer new relationships* that have not been directly observed, thus enabling adaptive and flexible behaviour.

Second, new data samples can be used to provide *training data* to neural circuits, to *drive learning* in the absence of sensory input. This training signal may manifest as an internally generated prediction error signal, using the mismatch between: (i) latent state representations used to simulate sensory data in low levels of the hierarchy and (ii) *inferred* latent state representations, where the simulated sensory data are used as input for perceptual inference. Thus, in the absence of sensory data, the resulting prediction error signal can be used to update neural circuits, including the parameters of the generative model itself.

From a computational perspective, the Helmholtz machine [1] provides an intuitive example of how a generative model may be implemented in neural architecture. In brief, the Helmholtz machine is a hierarchical model that uses a recognition model to infer the latent causes of the observed sensory data. The recognition model allows the inversion of the internal world model, to enable inference of latent causes from sensory data. This recognition model is then coupled with a generative model capable of generating sensory data from latent causes represented by the internal world model.

In the Helmholtz machine, learning is achieved via a 'wake-sleep' algorithm [6] that separates opportunities for learning. During the wake phase, sensory data from the world drive the lowest layers and activity propagates up to higher layers, where the underlying cause is inferred using a feedforward approximation of the posterior. During the sleep phase, the generative model is used to sample latent variables (or priors), thus generating example data, which are first represented by the top layers before activity propagates to successively lower layers. Since the generative model represents the true underlying cause for the generated data (e.g. 'banana'), the target value of the hidden units is known, providing a means to train synaptic weights in lower layers of the hierarchy. In other words, the sleep phase uses self-generated training data to train the recognition model to perform rapid feedforward approximation of the posterior. Notably, the generative model allows the recognition model to invert the generative model for inputs distributed according to probability distributions in the generative model rather than distributions in real data. In this manner, the generative model calibrates perceptual inference. The Helmholtz machine is arguably the conceptual precursor for incorporating generative capacity in modern AI, including variational autoencoders (VAEs), where VAEs learn how to represent a set of probabilistic latent variables that can be sampled to generate new data. VAEs have been used to provide elegant computational accounts of the hippocampus as a generative model [7,8].

Here we explore how the hippocampus, a brain region important for memory, can be characterized as a generative model [7–18]. Our opinion piece aims to provide a unique synthesis of empirical data and computational theory. For example, we explore the distinct anatomical position of the hippocampus, which sits at an apex in the cortical hierarchy and receives converging inputs from disparate neural circuits. We suggest, therefore, that the hippocampus is uniquely positioned to execute the dual aspects of a generative model. Namely, during 'online' behaviour the hippocampus can contribute to *perceptual inference*, while in 'offline' periods of rest or sleep, the hippocampus has unique capacity to intrinsically generate spiking sequences that resemble *self-generated training data*. The synergy of these two aspects of a generative model may be unique to the hippocampus, yet the underlying computations operate in concert with other circuits. For example, perceptual inference is implemented via message passing up the cortical hierarchy, while spiking sequences that are intrinsically generated in hippocampus can propagate down the cortical hierarchy, to even predict activity in primary sensory cortices [19–21]. We highlight how these mechanisms interact with neuromodulatory systems, such as the dopaminergic system. Moreover, we explore how higher-order brain regions, such as entorhinal cortex and prefrontal cortex, may provide structure to both hippocampal representations of latent causes and spiking sequences generated within the hippocampus. We formulate a series of predictions that arise from this perspective (see Box 1). Taken together, our perspective provides a comprehensive characterization of the hippocampus as a generative model that formulates a natural extension to the more traditional view that the hippocampus represents a cognitive map [22–27].

Box 1. Predictions*1. Hippocampal theta sweeps reflect generative sampling*

Hippocampal spiking activity in late theta phases reflects generative sampling of future inferred latent states. This ‘online’ generative activity can rapidly predict upcoming behaviour, over a short-time horizon. Silencing the CA3 region of the hippocampus in late theta phases should impair behaviours that rely on short-term planning, as recurrent connectivity in CA3 should endow the CA1 region of the hippocampus with sequences that carry online generative capacity.

2. Internally generated prediction errors can be used to update the generative model during offline periods of rest/sleep

Internally generated prediction error signals are computed in the hippocampus during rest/sleep, in the absence of sensory input. These prediction error signals are generated using the mismatch between (a) hippocampal latent state representations used to simulate sensory data in low levels of the cortical hierarchy, and (b) *inferred* hippocampal latent state representations, where the simulated sensory data provide input that can be used for perceptual inference. These prediction error signals may be used to update neural circuits, including the parameters of the generative model itself.

3. Dopamine projections to hippocampus are necessary to infer new latent states in the hippocampus

A sufficiently large dopaminergic prediction error signal can drive formation of a new latent state representation in the hippocampus. Blocking dopaminergic projections to hippocampus (deriving from the ventral tegmental area (VTA) and/or locus coeruleus (LC)) will impair inference of new latent states. For example: in rapid extinction paradigms, this manipulation will reduce expression of renewal behaviour; following changes in context, the probability of global remapping will be reduced; on a new task, splitter cells may not emerge to represent hidden latent states.

4. Generative replay plays a causal role in assigning credit to previously neutral and/or non-contiguous stimuli

Generative replay sequences that start with representation of reward provide opportunity to assign credit to previously neutral or non-contiguous cues by ‘back-propagating’ value to other cues. Dopaminergic projections to hippocampus are predicted to be necessary for this mechanism. Therefore, blocking dopaminergic projections to hippocampus should impair credit assignment to neutral or non-contiguous cues, while sparing learning for directly reinforced associations.

5. Structural codes in medial entorhinal cortex (mEC) are iteratively learned from and applied to theta sweeps, and in turn constrain the structure of generative replay during Sharp-Wave Ripples (SWRs)

Low-dimensional representations in mEC can impose structure on hippocampal activity. These structural codes are predicted to be iteratively learned from and applied to hippocampal spiking sequences in theta, during active behaviour. In turn, mEC may constrain the structure of generative replay during offline SWRs. Silencing mEC during hippocampal SWRs should release this constraint, allowing hippocampal sequences to deviate from plausible and optimized future narratives.

6. Symptoms reported in psychosis can be explained by impaired segmentation of latent states

In psychosis, mechanisms that may contribute to the formation of delusional beliefs include: (a) hyperactivity in the hippocampus, leading to aberrant or ‘over-extended’ generative activity (in both theta and SWRs) that has potential to bind non-contiguous stimuli and/or misattribute credit; (b) impaired dopaminergic signalling, which under-segments latent state representations in hippocampus, and attributes excess salience to inferred latent states, accounting for persistent and strong convictions of delusional beliefs.

Overall, by conceptualizing the hippocampus as a generative model, we describe how the hippocampus contributes to broad aspects of higher-order cognition that extend beyond memory *per se*, to include state representation, decision-making, inference, imagination, planning, continual learning and the extraction of statistical regularities. Further, by examining how temporal sequences within the hippocampus provide a generative readout in time, we account for representations of latent *temporal* structure in hippocampus, both during abstract tasks [28] and in response to rich narratives [29]. Finally, we discuss how perturbations to the hippocampus may disturb the capacity of the generative model to support adaptive behaviour. Using psychosis as a case study, we suggest that perturbations to the hippocampal generative model can account for core clinical symptoms in psychosis, including hallucinations and delusions.

2. From cognitive map to generative model

The traditional view of the hippocampus is that it creates a cognitive map that can be used to navigate both physical and more abstract spaces. The term ‘cognitive map’ was first introduced by Tolman to explain behaviour in rodents. Namely, after being exposed to a spatial maze, rats were able to infer shortcuts to obtain rewards [22] and infer new routes when old routes were blocked [23]. To account for these behaviours, Tolman suggested that the mammalian brain represents a rich internal model of the world or cognitive map that provides a means to predict the consequence of actions, including those that involve novel trajectories.

The discovery of ‘place cells’ in the rodent hippocampus provided the first evidence for a neural implementation of the cognitive map. Place cells are pyramidal cells in the hippocampus that show elegant tuning to specific spatial locations [24]. Together with a rich menagerie of other spatial codes reported in the hippocampus and neighbouring brain regions, place cells provide a model of *where* an agent is in absolute space. The model serves a memory system by representing locations

in an organism's environment, their spatial relations and the existence of specific objects in specific places. This allows an agent to locate itself and navigate within a familiar environment without reference to any specific sensory input. O'Keefe & Nadel described this model as a 'cognitive map': a neural system that generates a model of absolute space [24]. Equivalent neural codes can be observed for non-spatial relationships [25,26], suggesting that the hippocampal cognitive map represents relationships that not only help us navigate the spatial environment but also more abstract relationships and knowledge [27].

Here, by exploring the broader anatomical position and function of the hippocampus, we seek to extend this view to suggest the hippocampus can be conceptualized as a *generative model*. As outlined above, the first key function of a generative model is to perform perceptual inference by inferring and representing the latent causes of sensory data that are not directly observed [30]. Below we examine theoretical and empirical evidence to suggest the hippocampus is ideally positioned to support perceptual inference on highly processed sensory input [7,16,18]. Within this view, the hippocampus does not merely represent a cognitive map or internal model (hereafter used interchangeably) built from past experience (or memory). Rather, the hippocampus represents the inversion (or approximate inversion) of the internal model to map observed sensory data to the latent (unobserved) causes or states (i.e. *perceptual inference*). In other words, the hippocampus represents the posterior as a latent cause that draws from both past experience *and* incoming sensory data. Within this framework, place cells may be interpreted as representing the unobserved location in the cognitive map that is inferred from sensory input, namely the outcome of latent state inference.

However, often the inferred state is more abstract than a spatial location. For example, the latent cause for a particular sensory experience may resemble a particular trajectory through an internal model [8] which may be highly abstract relative to our perceived structure of the world. It is difficult to conceptualize how such highly abstract latent causes should be represented. To a human observer of neural data, activity may be assumed to map onto fundamental axes of perceived structure, such as space and time. However, in the brain the relevant metrics for an internal model are those that can be utilized by downstream neural observers [31]. Therefore, when an agent traverses a linear track, the latent causes for the sensory input may appear to be represented as locations in space (via 'place cells') and can be plotted as a trajectory on a spatial map. When an agent traverses a more abstract experience, such as holding a conversation with a colleague, latent causes for the sensory input may be represented in a manner that is more difficult to interpret, simply because the low dimensional embedding of the representation does not readily map onto familiar dimensions of space and time.

The second key function of a generative model is to *generate new data samples*, by constructing data samples from the representation of latent causes. As discussed below, the hippocampus has unique capacity to intrinsically generate new data samples (i.e. 'replay'). While a cognitive map can be used to draw new data samples, generative models provide greater explanatory power by naturally accounting for the computational machinery needed to intrinsically generate new data. Thus, we explore both theoretical and empirical evidence to support the view that the hippocampus both represents latent causes and has capacity to generate new data samples. In doing so, we characterize the hippocampus as a generative model, which naturally extends and formalizes the notion of a cognitive map within a broader computational framework.

3. The hippocampus represents and generates new latent causes

In this section we define how the hippocampus contributes to inferring the latent and unobserved causes of sensory data ('*perceptual inference*'). By examining empirical data and theoretical frameworks, we propose that latent state inference is computed across the cortical hierarchy, where the unique contribution of the hippocampus and related circuits is to create and utilize representations of *new latent causes*. Notably, perceptual inference is not a feature of several popular computational accounts of hippocampal function, such as the successor representation [32,33], which otherwise explain a number of properties of place cell and grid cell activity.

Although empirical evidence for perceptual inference predominantly derives from visual cortices [34], a Bayesian mathematical framework can be applied to the entire cortical hierarchy. Therefore, at each level of the cortical hierarchy, descending predictions (or priors) that originate in high levels of the cortical hierarchy are integrated with ascending sensory inputs to infer the latent cause by estimating a *posterior*. Over time, as dynamic message-passing continues to ascend and descend the cortical hierarchy, the representation of a latent cause in lower levels of the cortical hierarchy may become increasingly abstract.

While it remains experimentally difficult to assess dynamic message-passing across the cortical hierarchy, both computational modelling and empirical evidence suggest representations of latent states can be found in higher-order brain regions, including the prefrontal cortex and orbitofrontal cortex, as discussed elsewhere [35–37]. Here, we focus on the contribution of hippocampus to latent state inference. First, causal manipulations in mice demonstrate that the hippocampus is necessary for inferring latent context [38]. Second, the phenomenon of hippocampal remapping suggests the hippocampus represents latent causes. Hippocampal remapping can be defined as the change in preferred spatial tuning or firing intensity of place cells between different environments. Therefore, at the population level, remapping involves a change in the map that tiles the environment. This change in the map can involve *global remapping* where the preferred spatial tuning of hippocampal place cells changes, or *rate remapping*, where the preferred spatial tuning remains the same but the firing rate changes [39].

The occurrence of rate versus global remapping appears to depend on both the nature of the environmental change and whether the animal infers that it is in a new context (latent cause). For example, global remapping is observed in response to substantial changes in sensory cues [40,41] that are sufficient to suggest a change in context represented as a new latent cause. In line with theoretical predictions [8], global remapping is not random but rather involves structured realignment of population activity [42,43]. This structured realignment may be determined by latent structural knowledge represented by grid cells in

entorhinal cortex [8]. Global remapping in the hippocampal population code can be interpreted as representing a different latent state.

Rate remapping, on the other hand, can be observed in response to more subtle changes in sensory cues. For example, Leutgeb *et al.* [39] recorded spiking activity in the hippocampus while rats were placed within a square environment with a specific wall colour. The colour of the walls was varied (either black or white), while leaving all other local and global cues in the environment intact. When the wall colour changed, the firing rate of pyramidal cells changed but their preferred spatial tuning remained the same (rate remapping) (figure 1a–c). This rate rather than global remapping in the hippocampal population code can be interpreted as maintaining the same latent cause (i.e. same environment) despite changes in sensory cues.

In addition to hippocampal remapping, so-called ‘splitter cells’ provide the hallmark characteristics of a neural representation of latent causes (figure 1d–f). While the traditional definition of a ‘place cell’ is a neuron that fires at a particular spatial location, ‘splitter cells’ transcend this definition by showing firing responses that are modulated by an animal’s recent choice history or upcoming goal [45–47]. For example, as an animal traverses the vertical arm of a continuous T-maze, the firing of a splitter cell will depend upon both the spatial location and whether the animal intends to take the left or right turn at the top of the vertical arm [48]. Therefore, the information represented by these cells extends beyond incoming sensory and motor input to include more abstract information about the recent past and/or upcoming choice. When the activity of neuronal ensembles is plotted in low-dimensional space, the neural activity appears to be ‘pulled apart’ or split by a hidden variable, despite the animal traversing the same physical location [49]. Therefore, activity in splitter cells does not necessarily correspond to observable states but instead reflects the agent’s beliefs about the underlying structure of the environment, given incoming sensory data [45]. Moreover, the activity of splitter cells can generalize across tasks with identical structure [49,50], providing further evidence to suggest that splitter cells may represent latent causes.

The phenomena of hippocampal remapping and splitter cells both provide evidence to support the view that the hippocampus represents unobserved latent causes of sensory input. These representations of latent causes may be inherited from another brain region or constructed anew in the hippocampus itself. By examining fear conditioning paradigms, we suggest that the hippocampus can play an active role in constructing representations of *new* latent causes. Here, an animal first undergoes fear conditioning (‘conditioning phase’, sound paired with a shock), before undergoing extinction (‘extinction phase’, sound without shock) until the animal no longer freezes in response to the sound. If the animal is later returned to the shock context, ‘renewal’ behaviour is typically observed, where the animal shows a fear response (freezing to sound without shock). Notably, renewal behaviour is observed when the extinction phase is introduced rapidly. The expression of renewal behaviour suggests that two separate latent causes (or posterior distributions) are used to explain the conditioning and extinction phase of the task. A mature and intact hippocampus appears necessary to infer new latent causes in response to rapid extinction, as both adult rats with hippocampal lesions and young rats with an intact hippocampus display a lack of renewal behaviour following fear conditioning with rapid extinction [51–53]. More nuanced reports of the effect of hippocampal lesion on renewal-like behaviour suggest that pronounced deficits are observed when both the extinction and conditioning occur within the same context [54]. This suggests that the hippocampus contributes to constructing latent state representations when there is sufficient ambiguity as to whether a new latent cause is appropriate. Alternative brain regions, such as prefrontal cortex and amygdala, may have capacity to compensate in certain paradigms [55].

In contrast to rapid extinction protocols, if extinction is introduced gradually, renewal is less pronounced. For example, following a gradual fear extinction paradigm in rats, Gershman *et al.* found no evidence for recovery of fear [44]. This suggests that gradual extinction promotes representation of a single latent cause for both the conditioning and extinction phases of the task. To provide a mechanistic account, one possibility is that, in a gradual extinction condition, the total prediction error across time is not sufficient to trigger inference of a new latent state. Rather, the latent state generated during the conditioning phase is slowly updated to gradually increase the likelihood of an extinction trial.

Dopaminergic modulation may play a key role in signalling the need for hippocampus to infer a new latent cause. When incoming sensory data are unexpected given prior beliefs, the resulting mismatch in expectation generates a dopaminergic prediction error signal. Redish *et al.* have proposed a theoretical model in which new latent causes are inferred from sensory data when a sufficiently large negative prediction error is encoded by dopamine, a process described as ‘state-splitting’ [10]. Here, acquisition of new learning and extinction are modelled as two interacting learning processes, which are driven by positive and negative reward prediction error signals, respectively. However, alternative theories suggest the magnitude, rather than the sign, of the prediction error determines ‘state-splitting’, where a small prediction error drives incremental learning, while a large prediction error generates a new state [13].

While empirical evidence is needed to arbitrate between these alternatives, the known anatomy supports the view that dopaminergic projections target hippocampus, providing a potential pathway to update and/or construct latent state representations. For example, dopaminergic fibres from the ventral tegmental area (VTA) preferentially innervate the output regions of the hippocampus (CA1 and subiculum) and causal manipulations demonstrate that these projections contribute to hippocampal-dependent memory performance [56]. In addition, projections from locus coeruleus (LC) to hippocampus are profuse and co-release dopamine with noradrenaline [57,58]. These LC projections to hippocampus signal a dopamine-dependent novelty effect which can mediate hippocampal learning [59]. VTA and LC dopaminergic innervation of hippocampus may therefore provide prediction error-like signals that drive inference of new latent states, with VTA conveying a signed reward prediction error signal and LC conveying a novelty signal. Conceptualizing the hippocampus as a generative model therefore leads to the prediction that dopaminergic innervation of the hippocampus is necessary to infer new latent states.

Theoretical and empirical work suggests that generative capacity within the hippocampus is also influenced by interactions between hippocampus and other higher-order brain regions, such as the medial entorhinal cortex (mEC). More specifically,

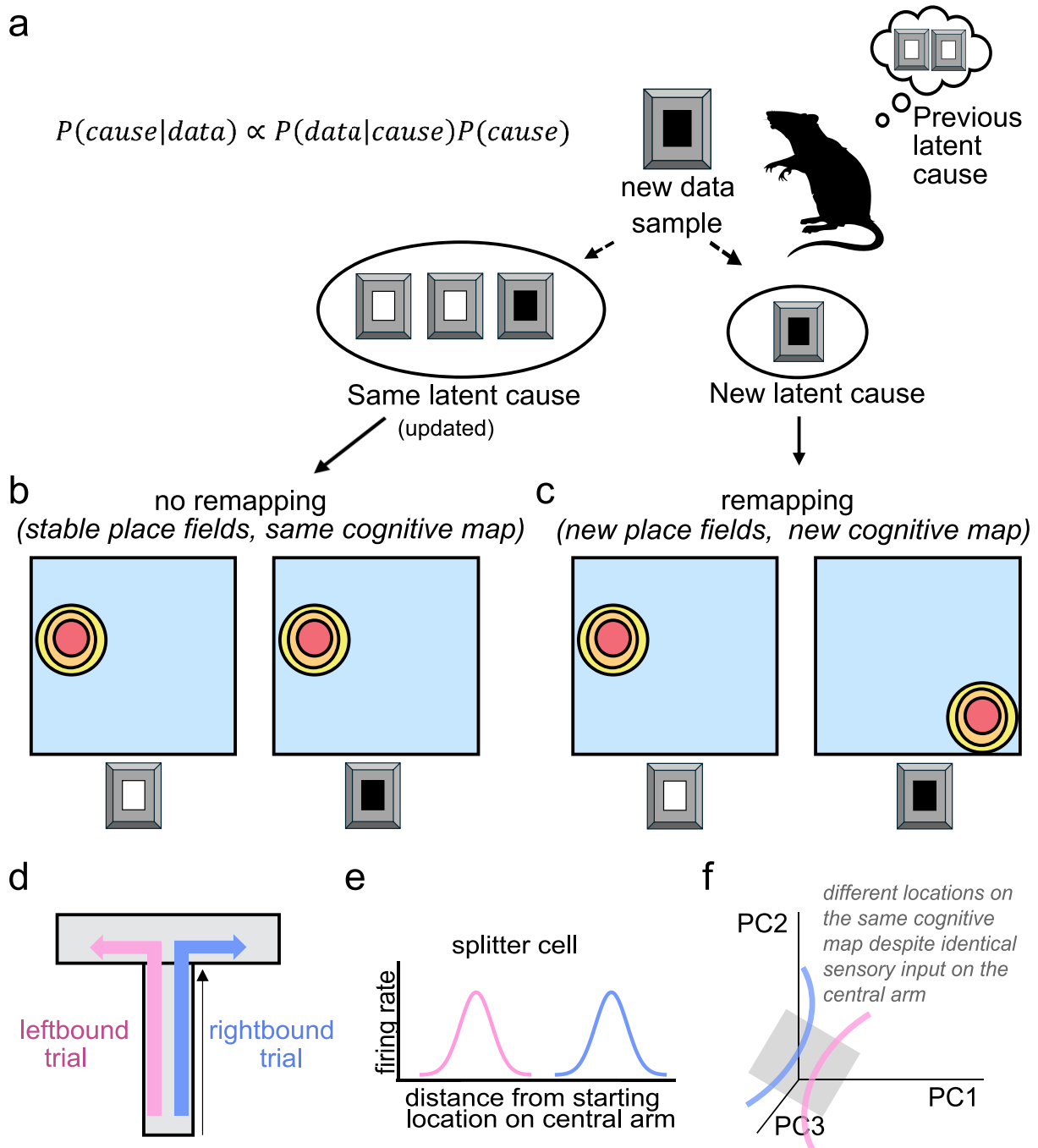


Figure 1. The hippocampus represents the posterior of latent state inference. (a–c) In conjunction with other brain regions, the probability of the current latent cause given incoming data is computed (a) (illustration adapted from [44]). If the incoming data are attributed to the same latent cause (b), the hippocampal preferred spatial tuning of place cells remains the same (no remapping, or ‘rate’ remapping). If the incoming data are attributed to a different latent cause (c), ‘global’ hippocampal remapping occurs to represent the new latent cause, which resembles a new cognitive map. (d) Example of a T-maze, where animals perform an alternation task (run to left arm, right arm, left arm) to obtain reward. (e) Schematic showing an example splitter cell in the hippocampus which represents different locations on the central arm of the T-maze depending on whether the upcoming trial is a left- or right-bound trial. (f) Schematic showing that when projected into a low-dimensional space, the activity of splitter cells reflects different locations within a cognitive map which may reflect a representation of two distinct latent causes, despite identical sensory input; (e,f) adapted from [45].

mEC is attributed with constraining the relevant set of possible latent causes (or posteriors) represented by hippocampus. Firing activity in pyramidal neurons in the mEC tiles both physical and abstract spaces using a grid-like lattice [25,60–62]. Mathematically, this grid-like code can be formulated as the low-dimensional embedding of activity in hippocampal pyramidal cells. Specifically, for place cells that represent location within a two-dimensional environment, grid-cell-like firing can be obtained from the eigenvectors of the covariance of the two-dimensional place cells, using the output of a principal component analysis (PCA) with non-negative weights [33,63]. The eigen code in mEC therefore provides an efficient means to capture the variance or structure of activity in the hippocampal population [27]. Neuroimaging studies in humans and computational models both provide evidence to suggest that the mEC provides an explicit representation of structure [8,18,28,33,64,65]. An interesting prediction that arises from this interaction between mEC and hippocampus is that the structural constraint imposed by mEC invokes non-random hippocampal remapping [8,15].

Insight into the mechanism by which mEC structural codes are generated and used can be gained by examining anatomical connections and causal relationships between the mEC and hippocampus. The CA1 region, which serves as a primary output region of the hippocampus, receives inputs from two alternative pathways: the trisynaptic path (via EC → dentate gyrus → CA3 → CA1) and the monosynaptic path (via EC → CA1). When considering the monosynaptic path, deep layers of EC (layer V/VI) receive direct projections from CA1, while superficial layers of EC (layer III) send direct projections back to CA1. Grid cells are primarily found in the superficial layers of mEC, providing opportunity to constrain CA1 activity using a structural code. In the temporal domain, this structural constraint on CA1 is reflected in theta phase precession. Bilateral inhibition of mEC disrupts theta in CA1 [66] and mEC lesions disrupt CA1 phase precession even during periods of stable spatial firing [67]. Together these observations suggest that mEC can both constrain the possible set of latent causes represented by CA1 and influence the precise temporal structure of these CA1 representations during active behaviour.

In addition to mEC, other entorhinal subregions may play an important role. For example, the lateral entorhinal cortex (IEC) projects to dopaminergic regions (VTA and substantia nigra, [68]) and, in turn, receives strong dopaminergic inputs from these regions [69]. Dopaminergic prediction errors across IEC and CA1 may therefore contribute to switches in the preferred set of latent causes. At a physiological level this may manifest as a phase shift in the grid cells (akin to grid cell remapping [70]). The interaction between CA1, IEC and dopaminergic prediction error signals in midbrain may therefore determine both whether a new latent state is represented in hippocampus, and which set of possible latent states can be represented by hippocampus.

4. The hippocampus generates new data samples

In addition to perceptual inference, generative models have capacity to *generate new data samples* (figure 2). These new data samples can be directly (re)constructed from the representation of latent causes. The capacity to generate new data samples provides an efficient computational solution to many aspects of higher-order cognition, such as memory consolidation, continual learning, the prediction of future observations and inference of new relationships. Here, we examine firing sequences reported in the hippocampus that reflect this aspect of generative activity.

First, we consider firing sequences that occur in the hippocampus during theta oscillations. During online periods, when animals or humans engage with a task, hippocampal pyramidal neurons fire at systematically earlier phases of the theta cycle (4–12 Hz in rodents, 2–8 Hz in humans) as the animal [73–78] or human [79–82] traverses through the cell's place field (figure 2a). This phenomenon is known as *theta phase precession*. While phase precession was first observed across place cells, the phenomenon has since been reported for more abstract, non-spatial sequences [26]. At the population level, coordinated phase precession across ensembles can give rise to theta sequences, which can be described as a *theta sweep*: a temporal ordering of neural activity that spans past, present and potential future locations within a single theta cycle. On a linear track and in a T-maze environment, forward-directed theta sweeps are observed [78,83,84], while in an open field, left–right alternating sweeps can be observed [85]. We note that phase precession can occur in individual neurons in the absence of coherent population-level theta, reflecting an intrinsic property of place cells [74]. However, while phase precession can be observed in individual neurons without prior experience, ensemble-level theta sequences develop rapidly with exposure to a new spatial maze [86], becoming stronger with experience [87]. Thus, theta sweeps, rather than phase precession intrinsic to individual neurons, may reflect rapid exploration of possible future trajectories to predict upcoming behaviour over a short-time horizon [83,88,89], although a reliable relationship between theta sweeps and behaviour is not always observed [70]. In addition to representing the immediate future, theta sweeps are also reported to represent counterfactual trajectories [84] and distal spatial locations [90], suggesting that reconstructed data samples have the potential to be non-local.

Theta phase precession therefore provides a candidate generative mechanism where late theta phases reflect new data sampled from latent causes. Examining the anatomy of the hippocampus adds nuance to this view. Along the trisynaptic pathway (via EC → dentate gyrus → CA3 → CA1), recurrent connectivity within CA3 has traditionally been considered to support memory recall or pattern completion. Here, we suggest that recurrent connectivity within CA3 endows CA1 with capacity to receive internally generated data samples that are reconstructed from an internal world model. By integrating inputs from this trisynaptic pathway with those of the monosynaptic pathway (via EC → CA1), CA1 has opportunity to compute a prediction error signal between generative activity and incoming sensory data (from EC) [91]. Empirical data support this view, with hippocampal blood oxygen level-dependent (BOLD) signal in humans covarying with the magnitude of prediction errors [92–95].

A second candidate firing pattern for generative hippocampal activity occurs during sharp-wave ripples (SWRs): high-frequency oscillations that dominate the hippocampal local field potential during offline periods of rest and sleep. The consensus view is that the genesis of a sharp-wave occurs within the highly recurrent CA3 subregion of the hippocampus, which is endowed with the high-frequency ripple generated by fast-spiking interneurons in CA1 [96]. During SWRs, firing activity in the hippocampus appears to 'replay' past experience on a temporally compressed scale that is concomitant with the induction of synaptic plasticity [90,97]. Selective disruption of hippocampal SWRs impairs learning and memory, demonstrating a causal role between SWR replay and memory consolidation [98–100]. In AI, the computational advantages of replay can be demonstrated in a continual learning setting. Continual learning is the ability to demonstrate new learning, without rapid forgetting of what was learnt before. Therefore, when a computational model has to learn multiple tasks, replay of previous experience ('experience replay') [101] can help protect the model from catastrophic forgetting, such that learning a new task does not necessarily interfere with performance on an older task. In the brain, firing sequences during SWRs may therefore be conceptualized as providing a training signal that consolidates new learning while protecting against memory interference. Indeed, while SWRs are intrinsically generated within the hippocampus, the SWR and sequences generated therein can propagate across

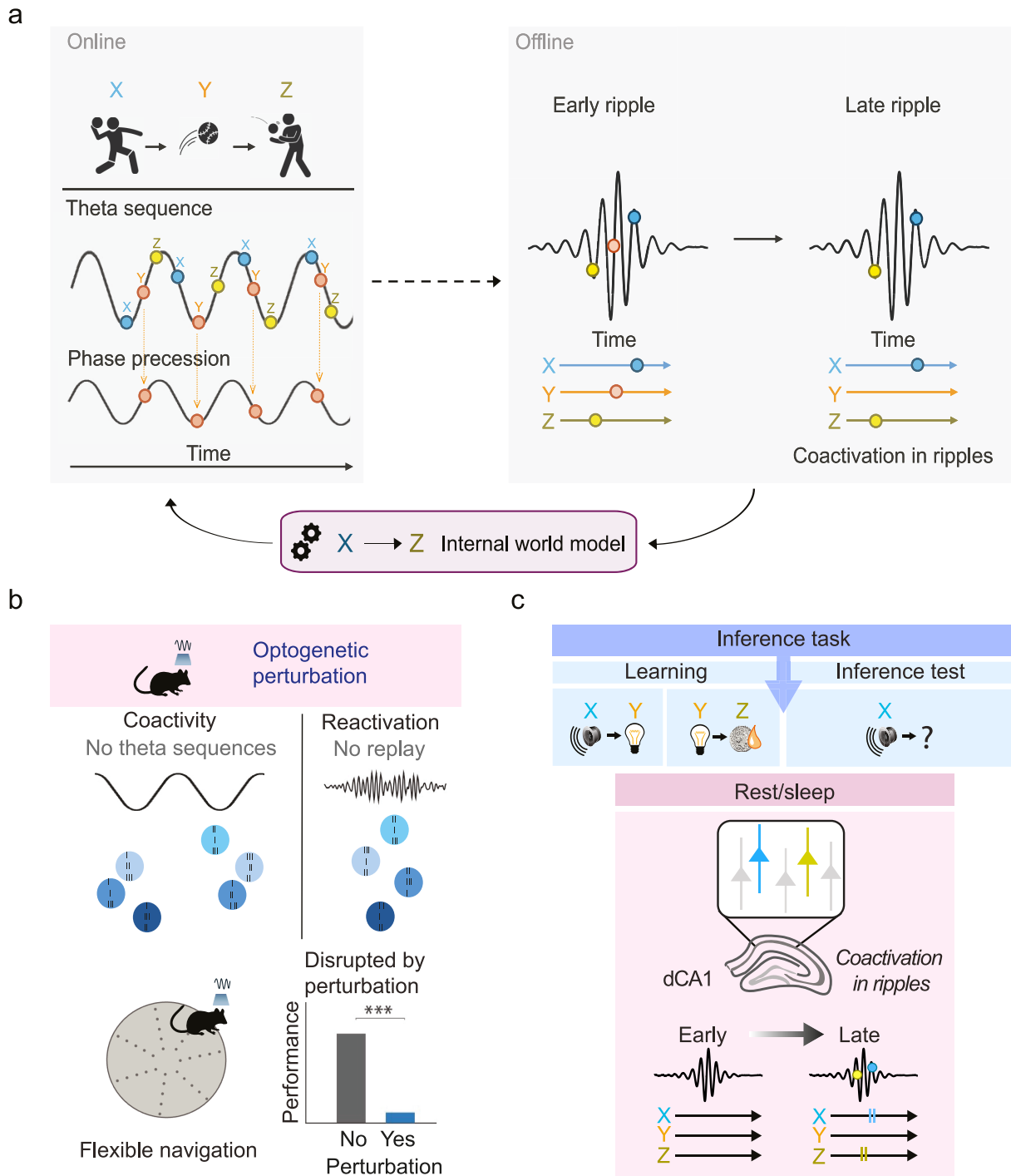


Figure 2. Hippocampal generative sequences in theta and sharp-wave ripples (SWRs). (a) *Left:* (Online) Schematic illustrating how theta phase precession organizes sequential elements within individual theta cycles. Each sequence element (X, Y, Z ; indicated by both colour and letter labels) is represented by coloured dots aligned to successive theta phases, producing a theta sweep from past (X) through present (Y) to future (Z). In this example, X corresponds to seeing a ball being thrown, Y to tracking the ball's trajectory, and Z to seeing the ball caught. *Right:* (Offline) Schematic showing generative replay during hippocampal SWRs in periods of rest and sleep. Early SWRs replay sequences of events (X, Y, Z). Over time, in late ripples, coactivation of X and Z is observed, without intermediary Y . Generative sequences across theta and SWRs may influence one another, such that plasticity from theta phase precession constrains the content of generative replay, while generative replay may provide a training signal to update the generative model. In turn, this influences theta phase precession upon re-exposure. (b) Figure adapted with permission from Liu *et al.* [71]. The temporal order of theta sequences in rats is disturbed using an optogenetic manipulation during ongoing spatial navigation. This manipulation disturbs the precise order of SWR firing sequences in subsequent periods of rest, suggesting that the content of theta sequences influences the content of generative replay. *** indicates $p < 0.001$. (c) Figure adapted with permission from Barron *et al.* [72]. In mice, shortcuts in the generative world model emerge over the course of hippocampal SWRs, with late SWRs skipping directly learnt associations to instead represent inferred relationships. dCA1 refers to the dorsal CA1 region of the hippocampus.

the cortical hierarchy, to potentially drive consolidation in distributed circuits that extend beyond the hippocampus [100,102]. When propagating across circuits, the temporal structure of the firing sequences must be preserved despite varying intrinsic timescales, to construct and propagate rich narrative structure [29]. Taken together, both prediction error signals and replay that are initially local to the hippocampus may reach all the way out to sensory cortex, providing an internally generated *training signal* to embed a hierarchical generative model across cortex.

However, in addition to merely replaying past experience, evidence suggests replay is generative, with spiking activity during SWRs representing sequences that extend *beyond* direct experience. One possibility is that this *generative replay* provides a *signature of planning*. For example, when rats perform goal-directed spatial navigation in an open arena, SWR-related firing sequences predict immediate future behaviour [103]. Moreover, SWR-related firing sequences can depict evaluation of potential options at a decision point [104] and reflect traversals of an aversive path before an animal shows avoidance behaviour [105]. However, when the content of awake SWRs is directly related to both past and future behaviour, replay content is found to be decoupled from subsequent choice [106]. Instead, replay content represents previously rewarded locations and places that have not been visited recently [106]. This suggests that generative replay may not provide a signature of planning for immediate upcoming behaviour. Rather, as explored in numerous computational models and theoretical perspectives [107–111], generative replay may further protect against catastrophic forgetting [112] and reflect adaptive computations that train a generative model to optimize future behaviour, rather than directly dictating immediate choices. Indeed, replay has long been considered a mechanism for memory consolidation [100].

To optimize future behaviour, generative replay can provide virtual exploration of hypothetical trajectories to *maximize future reward*. For example, replay is implicated in computational models, such as the successor representation [113,114], which use replay to rapidly propagate reward information and learn long-range predictive structures, to efficiently assign credit and learn multi-step relationships that extend beyond direct experience. Empirical data support this view. Spiking activity during SWRs can, for example, represent spatial trajectories that have never been experienced by an animal [105], future event locations [104], spatial trajectories to future goal locations [103] and shortcuts that support inference to a reward [72]. The direction of generative replay further suggests that sequences anchor around real and/or hypothetical rewards. For example, during waking behaviour, replay often proceeds in the forward direction to prospectively represent trajectories toward future goals [103,104], whereas receipt of reward is often accompanied by reverse replay [72,115,116]. This reverse replay provides a potential mechanism to back-propagate reward or assign credit to the start of a trajectory [103,115], or even to a non-contiguous cue that logically leads to reward [72] (figure 2c). This account may be extended to salience attribution to otherwise neutral cues, given evidence to suggest that novel, surprising or affective memories are preferentially replayed [56,98,117].

Causal manipulations in humans further support the idea that generative replay favours hypothetical trajectories to maximize future reward. These causal manipulations use a technique called targeted memory reactivation (TMR), which involves pairing sensory cues with specific memories during learning, before re-exposing participants to these cues during offline periods of sleep/rest [118–120]. This re-exposure biases firing sequences within hippocampal SWRs towards the cued associations [121], thus providing a tool to causally manipulate the content of replay. The behavioural consequences of TMR vary depending on when and how TMR is delivered. When TMR is applied as a contextual manipulation during awake rest, TMR can bias reactivation towards an entire set of memories without interfering with the endogenous replay sequences. Thus, contextual TMR can be used to enhance generative replay across stimuli or memories associated with the contextual cue. Intriguingly, following a contextual TMR manipulation, participants are subsequently more accurate at inferring the value of an outcome, while no effect is observed for behaviours that draw directly from previous experience [122,123]. This suggests that, rather than merely supporting memory consolidation or planning, firing sequences during SWRs can reflect the resampling of fictive sequences from a generative model [18]. These fictive sequences may provide an efficient means to explore *both real and hypothetical trajectories*.

For generative replay to be adaptive, the generated sequences must be constrained to a set of probable or plausible possibilities. In other words, the rules or grammar defining the structure of generative replay must be non-random. The empirical data discussed above provide some insight into these rules, suggesting that replay is shaped by behavioural relevance, which is influenced by reward and salience [100,102]. Computational modelling provides additional insight into the rules that govern the content of generative replay. Within a reinforcement learning framework, generative replay favours trajectories that update those internal representations (priors) with greatest potential to reduce future uncertainty and/or maximize reward [107,124]. When considering generative replay and structural learning in concert, EC is thought to impose structure on hippocampal firing sequences [8], to constrain generative replay to sequences that are probable and/or plausible. In other words, EC may define the low-dimensional embedding space (or manifold) through which hippocampal generative replay can traverse. By considering the sensory inputs received by EC, the function of EC can further be conceptualized as transforming egocentric sensory representations to allocentric latent variables represented by hippocampus [125–127].

This interaction between EC and hippocampus resembles that observed during theta (discussed above) where EC inputs to CA1 impose structure on theta sequences to ensure tractable latent state inference. This raises the intriguing possibility that structural codes in EC are iteratively learnt from and applied to both theta sweeps and, in turn, generative replay during SWRs. Indeed, meaningful variation across both theta and SWR sequences may be learnt by plasticity rules such as spike-time-dependent plasticity or behavioural timescale synaptic plasticity [128]. For example, within a theta sweep, future and past events are mapped onto pre- and postsynaptic activity, respectively, within a time window concomitant with plasticity. During SWRs, where the time window for plasticity is further compressed, plasticity can occur not only between neighbouring events in a sequence but also between entirely disparate events that were spatiotemporally discontinuous in the waking period. Optogenetic perturbation of theta sequences during spatial navigation in a novel maze disrupts the sequential structure in SWRs [71], suggesting that theta sequences themselves may mediate plasticity required for subsequent replay (figure 2a,b). To summarize, theta and SWR sequences therefore provide opportunity for plasticity within hippocampus, which can provide a training signal to cortical circuits, to embed and refine a hierarchical generative model and in turn place constraints on generative sequences.

Comparisons between the biological brain and AI further emphasize the computational significance of generative replay. Notably, AI agents are often trained extensively using supervised learning algorithms that train an agent to predict labelled

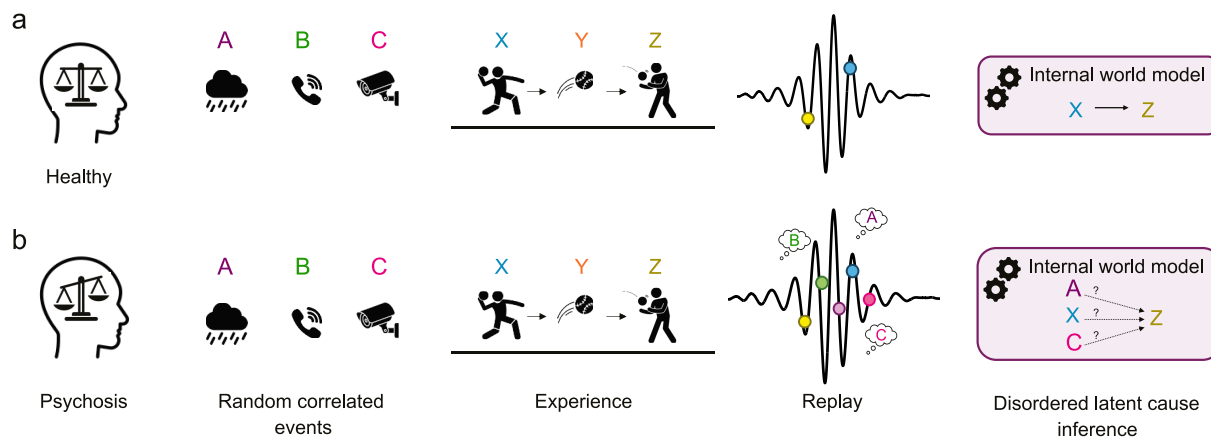


Figure 3. Perturbations in the hippocampal generative model explain core symptoms of psychosis. (a,b) Schematics comparing hippocampal generative processing in healthy controls (a) and psychosis (b). X corresponds to seeing a ball being thrown, Y to tracking the ball's trajectory, and Z to seeing the ball caught. A , B , C are random correlated events: A is rain, B is a phone call in the background, and C is a security camera in the vicinity. (a) Schematic for healthy controls. Random correlated events are not replayed in sharp-wave ripples (SWRs) during periods of rest or sleep. Latent causes are inferred correctly ($p(Z|X)$). (b) Schematic for people with psychosis. Owing to reduced effect of inhibitory interneurons, hippocampal hyperactivity allows random correlated events to be coactivated during SWR replay, leading to spurious sequences, disordered latent cause inference ($p(Z|A)$ or $p(Z|X)$ or $p(Z|C)$) and the formation of false beliefs. In this example they might infer that the ball is caught (Z) because someone is watching through the security camera (C) and not because it was thrown (X).

outputs. However, the brain rarely has access to explicit labels. Humans and animals are nevertheless able to rapidly and flexibly adapt to new environments. Applying self-supervised learning (SSL) to AI demonstrates that SSL can provide a means to efficiently explore real and hypothetical trajectories, to extract and learn an embedding space that captures meaningful factors of variation. Generative replay can provide the brain with a means to implement SSL, by replaying and traversing hypothetical trajectories, particularly after being exposed to a new experience.

5. Perturbations to the hippocampal generative model in psychiatric disorders

Given the computational reach of hippocampal generative activity, failures of this system may account for an array of different symptoms reported in neuropsychiatric disorders, including symptoms that extend beyond memory impairment. In this section, we use psychosis as a case study to explore how perturbations to the hippocampal generative model can account for core clinical symptoms, including hallucinations and delusions.

Psychosis is observed in several neuropsychiatric disorders and is considered a hallmark symptom of schizophrenia and bipolar disorder. Psychosis involves patients experiencing hallucinations and forming delusional beliefs. While behavioural assessments of these so-called 'positive symptoms' are used for diagnosis, positive symptoms are typically also accompanied by both negative and cognitive symptoms, including memory impairment [129]. Together these behavioural symptoms can be framed as impairments in perceptual and contextual inference, whereby individuals fail to appropriately segment experience into distinct latent causes [130–132]. In other words, patients have difficulty inferring *what* is causing *what* in the world. For example, delusions (e.g. 'they are watching me') may be attributed to under-segmenting latent causes, leading to failures in belief-updating despite changes in the underlying latent cause. This results in inappropriate generalization across contexts with poor separation between self and other, internal and external etc. Hallucinations (e.g. 'I can hear a voice') may similarly be attributed to inappropriate assignment of latent causes. Closely related constructs can be probed in psychosis using behavioural paradigms such as reversal learning paradigms [133], volatility estimation, and belief-updating under uncertainty [134]. However, here we highlight the need for more sensitive behavioural paradigms that explicitly aim to quantify disruption in latent state inference.

Framing behavioural symptoms in psychosis as an impairment in segmenting distinct latent causes provides some key predictions for our understanding of the underlying physiological mechanisms and has potential to guide development of novel treatment strategies. Current treatment for psychosis involves antipsychotic drugs that were largely discovered through serendipity. These drugs target excess dopamine via dopamine antagonists that focus predominantly on D2 receptors. While antipsychotic drugs can effectively reduce positive symptoms, such as hallucinations or delusions, treatment remains ineffective in around a third of patients and these drugs show no or limited effect on cognitive symptoms such as memory impairments [135]. This suggests that dopaminergic dysfunction alone is insufficient to account for the full spectrum of symptoms reported in psychosis. A broader framework, centred on the interaction between dopaminergic circuits and other brain regions, including the hippocampal generative model, may provide a more comprehensive account. One possibility is that a disordered interaction between hippocampus and dopaminergic circuits can account for impaired segmenting of distinct latent causes and perturbed updating of latent causes.

In individuals with psychosis, growing evidence suggests both structural and functional abnormalities in the hippocampus [136,137]. Structural changes in the hippocampus, such as volume decrease, are present in people at genetic risk for schizophrenia [138], and genome-wide association studies (GWAS) show a genetic overlap between the risk for schizophrenia and lower

hippocampal volume [139]. In addition to these structural changes, functional changes include hippocampal hyperactivity, which is thought to be attributed to increased neural firing in excitatory pyramidal cells in the hippocampus [140–145]. This increased neural firing in excitatory pyramidal cells may in turn be attributed to pre- and postsynaptic abnormalities in γ -aminobutyric acid (GABA)ergic interneurons, particularly parvalbumin-positive (PV+) GABAergic interneurons, which show hypofunction in psychosis [146,147]. Additional GABAergic markers suggest abnormalities in both patients with psychosis and genetic risk carriers [145,148–150], indicating reduced cortical inhibitory firing and reduced hippocampal GABA concentrations. Abnormalities in hippocampal PV+ GABAergic inhibitory interneurons may reduce the selectivity of pyramidal cells, analogous to reduced selectivity of place fields observed in the developing hippocampus prior to PV+ interneuron maturation [151]. Across the population, reduced selectivity of place fields may manifest as disruption in remapping. Therefore, the hippocampus may fail to remap when new latent causes should be inferred, reflecting an impairment in segmenting latent causes.

This aberrant remapping due to reduced selectivity of hippocampal pyramidal neurons may be further exacerbated by the effect of hippocampal hyperactivity on downstream brain areas. For example, hippocampal hyperactivity may drive excess release of dopamine in dopaminergic VTA [152,153]. Excess release of dopamine may perturb prediction error signalling to hippocampus, further disrupting the mechanisms that support the formation of new latent states or transitions between existing states. Moreover, excess dopamine may assign inappropriate salience to inferred latent states in hippocampus, contributing to the persistence and strong conviction of hallucinations and delusional beliefs in psychosis.

Hippocampal hyperactivity may also account for changes in fine-tuned temporal sequences that support generative activity. Reduced theta power and reduced theta coherence are reported in both animal models of psychosis and people with lived experience of psychosis [154–158]. Gamma oscillations are often nested within theta cycles, forming theta–gamma phase–amplitude coupling that organizes hippocampal activity across timescales [159]. A reduction in the power of both gamma-band oscillations and theta–gamma phase amplitude coupling can be observed in psychosis, from people at clinical high risk to those with chronic schizophrenia [147,160–163]. In people at genetic risk of psychosis, reduced power of theta and gamma oscillation appears to result from a lack of age-expected maturation during adolescence, in agreement with a neurodevelopmental model of schizophrenia [164]. Consistent with a late maturation of PV+ interneurons [165], these changes in major oscillatory markers may be the direct consequence of PV+ hypofunction and associated hippocampal hyperactivity. Indeed, PV+ interneurons are reported to mediate phase coupling of spiking activity in hippocampal theta oscillations [166] and control temporal dynamics within gamma oscillations [167]. In psychosis, disturbances in hippocampal PV+ interneurons may therefore disrupt the precise timing of hippocampal theta sweeps and resulting predictions of upcoming sensory experience.

A similar logic can be applied to SWR firing sequences where gross changes in SWR power are observed in psychosis. For example, in mouse models of schizophrenia the rate of SWRs is elevated [168,169], while replay in patients with psychosis is less faithful to direct experience [170,171]. These reports suggest that hippocampal hyperactivity promotes spurious firing patterns that allow aberrant or ‘over-extended’ sequences. These aberrant sequences may account for core symptoms in psychosis (figure 3). For example, the interaction between aberrant sequences and excess dopamine may provide a mechanism to misattribute salience or value to previously neutral stimuli [172,173]. The interaction between hippocampal hyperactivity and dopaminergic circuits during SWRs may therefore provide a mechanism to *misattribute credit or value* in a manner that produces highly salient and rigid spurious latent causes. Taken together, failures in the physiological implementation of the hippocampal generative model may account for core symptoms in psychosis, providing testable prediction for the mechanistic basis of the disorder.

6. Manipulating the hippocampal generative model

Recent advances in neuroscience and non-invasive brain stimulation now allow causal manipulation of activity in the hippocampus, providing opportunity to experimentally manipulate generative activity. As discussed above, TMR provides a well established method to selectively bias the content of replay during periods of rest and sleep. Beyond sensory cueing, emerging non-invasive brain stimulation techniques permit direct perturbation of hippocampal dynamics. Transcranial focused ultrasound stimulation (TUS) can non-invasively target deep-brain structures, including the hippocampus, with unprecedented spatial precision [174]. Preclinical studies indicate that TUS can modulate hippocampal oscillatory coupling, including theta–gamma phase–amplitude interactions, to alter behavioural measures of both encoding and memory retrieval [175]. In humans, low-intensity, focal TUS can causally modulate hippocampal activity and selectively disrupt participants' ability to perform inference [176], establishing a powerful approach to further probe how the human hippocampus contributes to a generative model. Similarly, temporal interference (TI) stimulation can selectively modulate activity in deep structures, such as the hippocampus, by generating interference envelopes at behaviourally relevant frequencies (e.g. theta or ripple frequency range) through the superposition of multiple high-frequency fields [177]. Experimental work demonstrates that temporally patterned interference stimulation can entrain hippocampal oscillations and modulate memory-guided behaviour [178,179], providing a means to modulate generative replay. Together, these approaches offer promising opportunities to further characterize the hippocampus as a generative model and establish tools to manipulate the hippocampal generative model in patient groups.

7. Concluding remarks

Conceptualizing the hippocampus as a generative model offers a unifying framework for understanding the multifaceted role of the hippocampus in cognition. Beyond serving memory or representing a cognitive map, we propose that the hippocampus engages in perceptual inference by representing latent causes for sensory experience. Using these latent causes, the hippocampus has the capacity to then generate sequences that simulate possible futures, during both theta sequences and SWRs. This dual capacity to perform perceptual inference and generate new data samples enables the hippocampus to support a wide range of cognitive functions, including memory, prediction, imagination, planning, inferential choice and generalization. We further propose that perturbations to these generative processes may account for core symptoms reported in psychosis where internally generated experiences are mistaken for reality. The predictions that arise from this perspective are summarised in Box 1. We conclude that viewing the hippocampus as a generative model can explain how the hippocampus contributes to higher-order cognition while also offering a powerful framework to identify physiological mechanisms that underlie behavioural symptoms in psychiatric disorders.

Data accessibility. This article has no additional data.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. S.N.: conceptualization, project administration, writing—original draft, writing—review and editing; L.G.: conceptualization, project administration, writing—original draft, writing—review and editing; X.P.: conceptualization, project administration, writing—original draft, writing—review and editing; V.M.: conceptualization, project administration, writing—original draft, writing—review and editing; H.C.B.: conceptualization, project administration, resources, supervision, validation, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed herein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This research was funded by UKRI (MR/W008939/1 to H.C.B.), MRC (MR/W01971X/1, supporting L.G. and X.P.), a Clarendon Fund Scholarship (to S.N.) and an SNSF Postdoctoral Mobility fellowship (P500PM_217669 to V.M.). The Medical Research Council Centre of Research Excellence in Restorative Neural Dynamics is supported by the MRC (UKRI 936). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

- Dayan P, Hinton GE, Neal RM, Zemel RS. 1995 The Helmholtz machine. *Neural Comput.* **7**, 889–904. (doi:10.1162/neco.1995.7.5.889)
- Friston K. 2010 The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138. (doi:10.1038/nrn2787)
- Gershman SJ. 2019 The generative adversarial brain. *Front. Artif. Intell.* **2**, 18. (doi:10.3389/frai.2019.00018)
- Knill DC, Richards W(eds). 1996 *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press. (doi:10.1017/CB09780511984037)
- Doya K, Ishii S, Pouget A, Rao RPN(eds). 2011 *Bayesian brain: probabilistic approaches to neural coding*. Cambridge, MA: MIT Press.
- Hinton GE, Dayan P, Frey BJ, Neal RM. 1995 The 'wake-sleep' algorithm for unsupervised neural networks. *Science* **268**, 1158–1161. (doi:10.1126/science.7761831)
- Spens E, Burgess N. 2024 A generative model of memory construction and consolidation. *Nat. Hum. Behav.* **8**, 526–543. (doi:10.1038/s41562-023-01799-z)
- Whittington JCR, Muller TH, Mark S, Chen G, Barry C, Burgess N, Behrens TEJ. 2020 The Tolman-Eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. *Cell* **183**, 1249–1263. (doi:10.1016/j.cell.2020.10.024)
- Fuhs MC, Touretzky DS. 2007 Context learning in the rodent hippocampus. *Neural Comput.* **19**, 3173–3215. (doi:10.1162/neco.2007.19.12.3173)
- Redish AD, Jensen S, Johnson A, Kurth-Nelson Z. 2007 Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychol. Rev.* **114**, 784–805. (doi:10.1037/0033-295X.114.3.784)
- Gershman SJ, Niv Y. 2010 Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* **20**, 251–256. (doi:10.1016/j.conb.2010.02.008)
- Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. 2014 Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279. (doi:10.1016/j.neuron.2013.11.005)
- Niv Y. 2019 Learning task-state representations. *Nat. Neurosci.* **22**, 1544–1553. (doi:10.1038/s41593-019-0470-8)
- Sanders H, Wilson MA, Gershman SJ. 2020 Hippocampal remapping as hidden state inference. *eLife* **9**, e51140. (doi:10.7554/eLife.51140)
- Whittington JCR, McCaffary D, Bakermans JJW, Behrens TEJ. 2022 How to build a cognitive map. *Nat. Neurosci.* **25**, 1257–1272. (doi:10.1038/s41593-022-01153-y)
- George TM, Barry C, Stachenfeld K, Clopath C, Fukai T. 2023 A generative model of the hippocampal formation trained with theta driven local learning rules. *bioRxiv* 2023.12.12.571268. (doi:10.1101/2023.12.12.571268)
- Pezzulo G, Kemere C, van der Meer MAA. 2017 Internally generated hippocampal sequences as a vantage point to probe future-oriented cognition. *Ann. N. Y. Acad. Sci.* **1396**, 144–165. (doi:10.1111/nyas.13329)
- Stoianov I, Maisto D, Pezzulo G. 2022 The hippocampal formation as a hierarchical generative model supporting generative replay and continual learning. *Prog. Neurobiol.* **217**, 102329. (doi:10.1016/j.pneurobio.2022.102329)
- Sirota A, Csicsvari J, Buhl R, Buzsáki G. 2003 Communication between neocortex and hippocampus during sleep in rodents. *Proc. Natl Acad. Sci. USA* **100**, 2065–2069. (doi:10.1073/pnas.0437938100)
- Ji D, Wilson MA. 2007 Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat. Neurosci.* **10**, 100–107. (doi:10.1038/nn1825)
- Rothschild G, Eban E, Frank LM. 2017 A cortical–hippocampal–cortical loop of information processing during memory consolidation. *Nat. Neurosci.* **20**, 251–259. (doi:10.1038/nn.4457)
- Tolman EC, Honzik CH. 1930 *Introduction and removal of reward, and maze performance in rats*, pp. 257–275. Berkeley, CA: University of California Press.
- Tolman EC. 1948 Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208. (doi:10.1037/h0061626)
- O'Keefe J, Nadel L. 1978 *The hippocampus as a cognitive map*. Oxford, UK: Oxford University Press. See <http://www.cognitivemap.net>.
- Constantinescu AO, O'Reilly JX, Behrens TEJ. 2016 Organizing conceptual knowledge in humans with a gridlike code. *Science* **352**, 1464–1468. (doi:10.1126/science.aaf0941)
- Aronov D, Nevers R, Tank DW. 2017 Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature* **543**, 719–722. (doi:10.1038/nature21692)

27. Behrens TEJ, Muller TH, Whittington JCR, Mark S, Baram AB, Stachenfeld KL, Kurth-Nelson Z. 2018 What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* **100**, 490–509. (doi:10.1016/j.neuron.2018.10.002)
28. Schapiro AC, Turk-Browne NB, Norman KA, Botvinick MM. 2016 Statistical learning of temporal community structure in the hippocampus. *Hippocampus* **26**, 3–8. (doi:10.1002/hipo.22523)
29. Baldassano C, Chen J, Zadbood A, Pillow JW, Hasson U, Norman KA. 2017 Discovering event structure in continuous narrative perception and memory. *Neuron* **95**, 709–721. (doi:10.1016/j.neuron.2017.06.041)
30. Kemp C, Tenenbaum JB. 2008 The discovery of structural form. *Proc. Natl Acad. Sci. USA* **105**, 10687–10692. (doi:10.1073/pnas.0802631105)
31. Buzsáki G, Tingley D. 2018 Space and time: the hippocampus as a sequence generator. *Trends Cogn. Sci.* **22**, 853–869. (doi:10.1016/j.tics.2018.07.006)
32. Dayan P. 1993 Improving generalization for temporal difference learning: the successor representation. *Neural Comput.* **5**, 613–624. (doi:10.1162/neco.1993.5.4.613)
33. Stachenfeld KL, Botvinick MM, Gershman SJ. 2017 The hippocampus as a predictive map. *Nat. Neurosci.* **20**, 1643–1653. (doi:10.1038/nn.4650)
34. Lee TS, Mumford D. 2003 Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A* **20**, 1434–1448. (doi:10.1364/JOSA.20.001434)
35. Donoso M, Collins AGE, Koehlin E. 2014 Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science* **344**, 1481–1486. (doi:10.1126/science.1252254)
36. Moneta N, Grossman S, Schuck NW. 2024 Representational spaces in orbitofrontal and ventromedial prefrontal cortex: task states, values, and beyond. *Trends Neurosci.* **47**, 1055–1069. (doi:10.1016/j.tins.2024.10.005)
37. Ma F, Lin H, Zhou J. 2025 Prediction, inference, and generalization in orbitofrontal cortex. *Curr. Biol.* **35**, R266–R272. (doi:10.1016/j.cub.2025.02.021)
38. Mishchanchuk K, Gregoriou G, Qü A, Kastler A, Huys QJM, Wilbrecht L, MacAskill AF. 2024 Hidden state inference requires abstract contextual representations in the ventral hippocampus. *Science* **386**, 926–932. (doi:10.1126/science.adq5874)
39. Leutgeb S, Leutgeb JK, Barnes CA, Moser EI, McNaughton BL, Moser MB. 2005 Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science* **309**, 619–623. (doi:10.1126/science.1114037)
40. Anderson MI, Jeffery KJ. 2003 Heterogeneous modulation of place cell firing by changes in context. *J. Neurosci.* **23**, 8827–8835. (doi:10.1523/JNEUROSCI.23-26-08827.2003)
41. Bostock E, Muller RU, Kubie JL. 1991 Experience-dependent modifications of hippocampal place cell firing. *Hippocampus* **1**, 193–205. (doi:10.1002/hipo.450010207)
42. Fenton AA. 2024 Remapping revisited: how the hippocampus represents different spaces. *Nat. Rev. Neurosci.* **25**, 428–448. (doi:10.1038/s41583-024-00817-x)
43. Chen HT, Manning JR, van der Meer MAA. 2021 Between-subject prediction reveals a shared representational geometry in the rodent hippocampus. *Curr. Biol.* **31**, 4293–4304. (doi:10.1016/j.cub.2021.07.061)
44. Gershman SJ, Norman KA, Niv Y. 2015 Discovering latent causes in reinforcement learning. *Curr. Opin. Behav. Sci.* **5**, 43–50. (doi:10.1016/j.cobeha.2015.07.007)
45. Duvelle É, Grievies RM, van der Meer MAA. 2023 Temporal context and latent state inference in the hippocampal splitter signal. *eLife* **12**, e82357. (doi:10.7554/eLife.82357)
46. Ainge JA, Tamosiunaite M, Woergoetter F, Dudchenko PA. 2007 Hippocampal CA1 place cells encode intended destination on a maze with multiple choice points. *J. Neurosci.* **27**, 9769–9779. (doi:10.1523/JNEUROSCI.2011-07.2007)
47. Dudchenko PA, Wood ER. 2014 Splitter cells: hippocampal place cells whose firing is modulated by where the animal is going or where it has been. In *Space, time and memory in the hippocampal formation* (eds D Derdikman, JJ Knierim), pp. 253–272. Vienna, Austria: Springer. (doi:10.1007/978-3-7091-1292-2_10)
48. Wood ER, Dudchenko PA, Robitsek RJ, Eichenbaum H. 2000 Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron* **27**, 623–633. (doi:10.1016/s0896-6273(00)00071-4)
49. Takahashi S. 2013 Hierarchical organization of context in the hippocampal episodic code. *eLife* **2**, e00321. (doi:10.7554/eLife.00321)
50. Sun C, Yang W, Martin J, Tonegawa S. 2020 Hippocampal neurons represent events as transferable units of experience. *Nat. Neurosci.* **23**, 651–663. (doi:10.1038/s41593-020-0614-x)
51. Ji J, Maren S. 2005 Electrolytic lesions of the dorsal hippocampus disrupt renewal of conditional fear after extinction. *Learn. Mem.* **12**, 270–276. (doi:10.1101/lm.91705)
52. Yap CSL, Richardson R. 2007 Extinction in the developing rat: an examination of renewal effects. *Dev. Psychobiol.* **49**, 565–575. (doi:10.1002/dev.20244)
53. Corcoran KA, Maren S. 2001 Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. *J. Neurosci.* **21**, 1720–1726. (doi:10.1523/JNEUROSCI.21-05-01720.2001)
54. Frohardt RJ, Guarraci FA, Bouton ME. 2000 The effects of neurotoxic hippocampal lesions on two effects of context after fear extinction. *Behav. Neurosci.* **114**, 227–240. (doi:10.1037//0735-7044.114.2.227)
55. Bouton ME, Westbrook RF, Corcoran KA, Maren S. 2006 Contextual and temporal modulation of extinction: behavioral and biological mechanisms. *Biol. Psychiatry* **60**, 352–360. (doi:10.1016/j.biopsych.2005.12.015)
56. McNamara CG, Tejero-Cantero Á, Trouche S, Campo-Urriza N, Dupret D. 2014 Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. *Nat. Neurosci.* **17**, 1658–1660. (doi:10.1038/nn.3843)
57. Smith CC, Greene RW. 2012 CNS dopamine transmission mediated by noradrenergic innervation. *J. Neurosci.* **32**, 6072–6080. (doi:10.1523/JNEUROSCI.6486-11.2012)
58. Devoto P, Flore G. 2006 On the origin of cortical dopamine: is it a co-transmitter in noradrenergic neurons? *Curr. Neuropharmacol.* **4**, 115–125. (doi:10.2174/157015906776359559)
59. Takeuchi Tet al. 2016 Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature* **537**, 357–362. (doi:10.1038/nature19325)
60. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. 2005 Microstructure of a spatial map in the entorhinal cortex. *Nature* **436**, 801–806. (doi:10.1038/nature03721)
61. Bellmund JL, Deuker L, Navarro Schröder T, Doeller CF. 2016 Grid-cell representations in mental simulation. *eLife* **5**, e17089. (doi:10.7554/eLife.17089)
62. Bao X, Gjorgieva E, Shanahan LK, Howard JD, Kahnt T, Gottfried JA. 2019 Grid-like neural representations support olfactory navigation of a two-dimensional odor space. *Neuron* **102**, 1066–1075. (doi:10.1016/j.neuron.2019.03.034)
63. Dordek Y, Soudry D, Meir R, Derdikman D. 2016 Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. *eLife* **5**, e10094. (doi:10.7554/eLife.10094)
64. Garvert MM, Dolan RJ, Behrens TE. 2017 A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *eLife* **6**, e17086. (doi:10.7554/eLife.17086)
65. Bigus ER, Lee HW, Bowler JC, Shi J, Heys JG. 2024 Medial entorhinal cortex mediates learning of context-dependent interval timing behavior. *Nat. Neurosci.* **27**, 1587–1598. (doi:10.1038/s41593-024-01683-7)
66. Zutshi I, Buzsáki G. 2023 Hippocampal sharp-wave ripples and their spike assembly content are regulated by the medial entorhinal cortex. *Curr. Biol.* **33**, 3648–3659. (doi:10.1016/j.cub.2023.07.039)
67. Schlesiger MI, Cannova CC, Boubil BL, Hales JB, Mankin EA, Brandon MP, Leutgeb JK, Leibold C, Leutgeb S. 2015 The medial entorhinal cortex is necessary for temporal organization of hippocampal neuronal activity. *Nat. Neurosci.* **18**, 1123–1132. (doi:10.1038/nn.4056)
68. Agster KL, Pereira IT, Sadoris MP, Burwell RD. 2016 Subcortical connections of the perirhinal, postrhinal, and entorhinal cortices of the rat. II. Efferents. *Hippocampus* **26**, 1189–1212. (doi:10.1002/hipo.22603)

69. Lee JY *et al.* 2021 Dopamine facilitates associative memory encoding in the entorhinal cortex. *Nature* **598**, 321–326. (doi:10.1038/s41586-021-03948-8)
70. Fyhn M, Hafting T, Treves A, Moser MB, Moser EI. 2007 Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* **446**, 190–194. (doi:10.1038/nature05601)
71. Liu C, Todorova R, Tang W, Oliva A, Fernandez-Ruiz A. 2023 Associative and predictive hippocampal codes support memory-guided behaviors. *Science* **382**, eadi8237. (doi:10.1126/science.adi8237)
72. Barron H *et al.* 2020 Neuronal computation underlying inferential reasoning in humans and mice. *Cell* **183**, 228–243. (doi:10.1016/j.cell.2020.08.035)
73. Muller RU, Kubie JL. 1989 The firing of hippocampal place cells predicts the future position of freely moving rats. *J. Neurosci.* **9**, 4101–4110. (doi:10.1523/JNEUROSCI.09-12-04101.1989)
74. O'Keefe J, Recce ML. 1993 Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* **3**, 317–330. (doi:10.1002/hipo.450030307)
75. Jensen O, Lisman JE. 1996 Hippocampal CA3 region predicts memory sequences: accounting for the phase precession of place cells. *Learn. Mem.* **3**, 279–287. (doi:10.1101/lm.3.2-3.279)
76. Skaggs WE, McNaughton BL, Wilson MA, Barnes CA. 1996 Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* **6**, 149–172. (doi:10.1002/(SICI)1098-1063(1996)6:2<149::AID-HIPO6>3.0.CO;2-K)
77. Harris KD, Henze DA, Hirase H, Leinekugel X, Dragoi G, Czurkó A, Buzsáki G. 2002 Spike train dynamics predicts theta-related phase precession in hippocampal pyramidal cells. *Nature* **417**, 738–741. (doi:10.1038/nature00808)
78. Foster DJ, Wilson MA. 2007 Hippocampal theta sequences. *Hippocampus* **17**, 1093–1099. (doi:10.1002/hipo.20345)
79. Qasim SE, Fried I, Jacobs J. 2021 Phase precession in the human hippocampus and entorhinal cortex. *Cell* **184**, 3242–3255. (doi:10.1016/j.cell.2021.04.017)
80. Reddy *et al.* 2021 Theta-phase dependent neuronal coding during sequence learning in human single neurons. *Nat. Commun.* **12**, 4839. (doi:10.1038/s41467-021-25150-0)
81. Schonhaut DR, Rao AM, Ramayya AG, Solomon EA, Herweg NA, Fried I, Kahana MJ. 2024 MTL neurons phase-lock to human hippocampal theta. *eLife* **13**, e85753. (doi:10.7554/eLife.85753)
82. Guth TA, Brandt A, Reinacher PC, Schulze-Bonhage A, Jacobs J, Kunz L. 2025 Theta-phase locking of single neurons during human spatial memory. *Nat. Commun.* **16**, 7402. (doi:10.1038/s41467-025-62553-9)
83. Johnson A, Redish AD. 2007 Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* **27**, 12176–12189. (doi:10.1523/JNEUROSCI.3761-07.2007)
84. Kay K, Chung JE, Sosa M, Schor JS, Karlsson MP, Larkin MC, Liu DF, Frank LM. 2020 Constant sub-second cycling between representations of possible futures in the hippocampus. *Cell* **180**, 552–567. (doi:10.1016/j.cell.2020.01.014)
85. Vollan AZ, Gardner RJ, Moser MB, Moser EI. 2025 Left–right-alternating theta sweeps in entorhinal–hippocampal maps of space. *Nature* **639**, 995–1005. (doi:10.1038/s41586-024-08527-1)
86. Feng T, Silva D, Foster DJ. 2015 Dissociation between the experience-dependent development of hippocampal theta sequences and single-trial phase precession. *J. Neurosci.* **35**, 4890–4902. (doi:10.1523/JNEUROSCI.2614-14.2015)
87. Mehta MR, Lee AK, Wilson MA. 2002 Role of experience and oscillations in transforming a rate code into a temporal code. *Nature* **417**, 741–746. (doi:10.1038/nature00807)
88. Pastalkova E, Itskov V, Amarasingham A, Buzsáki G. 2008 Internally generated cell assembly sequences in the rat hippocampus. *Science* **321**, 1322–1327. (doi:10.1126/science.1159775)
89. Berners-Lee A, Wu X, Foster DJ. 2021 Prefrontal cortical neurons are selective for non-local hippocampal representations during replay and behavior. *J. Neurosci.* **41**, 5894–5908. (doi:10.1523/JNEUROSCI.1158-20.2021)
90. Yu JY, Frank LM. 2021 Prefrontal cortical activity predicts the occurrence of nonlocal hippocampal representations during spatial navigation. *PLoS Biol.* **19**, e3001393. (doi:10.1371/journal.pbio.3001393)
91. Lisman JE, Grace AA. 2005 The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* **46**, 703–713. (doi:10.1016/j.neuron.2005.05.002)
92. Aitken F, Kok P. 2022 Hippocampal representations switch from errors to predictions during acquisition of predictive associations. *Nat. Commun.* **13**, 3294. (doi:10.1038/s41467-022-31040-w)
93. Glitz L, Juechems K, Summerfield C, Garrett N. 2022 Model sharing in the human medial temporal lobe. *J. Neurosci.* **42**, 5410–5426. (doi:10.1523/JNEUROSCI.1978-21.2022)
94. Jang AI, Nassar MR, Dillon DG, Frank MJ. 2019 Positive reward prediction errors during decision-making strengthen memory encoding. *Nat. Hum. Behav.* **3**, 719–732. (doi:10.1038/s41562-019-0597-3)
95. Kumaran D, Maguire EA. 2006 An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS Biol.* **4**, e424. (doi:10.1371/journal.pbio.0040424)
96. Buzsáki G. 1986 Hippocampal sharp waves: their origin and significance. *Brain Res.* **398**, 242–252. (doi:10.1016/0006-8993(86)91483-6)
97. Foster DJ. 2017 Replay comes of age. *Annu. Rev. Neurosci.* **40**, 581–602. (doi:10.1146/annurev-neuro-072116-031538)
98. Girardeau G, Benchenane K, Wiener SI, Buzsáki G, Zugaro MB. 2009 Selective suppression of hippocampal ripples impairs spatial memory. *Nat. Neurosci.* **12**, 1222–1223. (doi:10.1038/nn.2384)
99. Jadhav SP, Kemere C, German PW, Frank LM. 2012 Awake hippocampal sharp-wave ripples support spatial memory. *Science* **336**, 1454–1458. (doi:10.1126/science.1217230)
100. Buzsáki G. 2015 Hippocampal sharp wave-ripple: a cognitive biomarker for episodic memory and planning. *Hippocampus* **25**, 1073–1188. (doi:10.1002/hipo.22488)
101. Rolnick D, Ahuja A, Schwarz J, Lillicrap T, Wayne G. 2019 Experience replay for continual learning. In *Advances in Neural Information Processing Systems 32 (NEURIPS 2019)* (eds HM Wallach, LH Larochelle, A Beygelzimer, F d'Alché-Buc, EB Fox, R Garnett), pp. 348–358. Red Hook, NY: Curran Associates.
102. Joo HR, Frank LM. 2018 The hippocampal sharp wave-ripple in memory retrieval for immediate use and consolidation. *Nat. Rev. Neurosci.* **19**, 744–757. (doi:10.1038/s41583-018-0077-1)
103. Pfeiffer BE, Foster DJ. 2013 Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**, 74–79. (doi:10.1038/nature12112)
104. Ólafsdóttir HF, Barry C, Saleem AB, Hassabis D, Spiers HJ. 2015 Hippocampal place cells construct reward related sequences through unexplored space. *eLife* **4**, e06063. (doi:10.7554/eLife.06063)
105. Wu CT, Haggerty D, Kemere C, Ji D. 2017 Hippocampal awake replay in fear memory retrieval. *Nat. Neurosci.* **20**, 571–580. (doi:10.1038/nn.4507)
106. Gillespie AK, Astudillo Maya DA, Denovellis EL, Liu DF, Kastner DB, Coulter ME, Roumis DK, Eden UT, Frank LM. 2021 Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron* **109**, 3149–3163. (doi:10.1016/j.neuron.2021.07.029)
107. Mattar MG, Daw ND. 2018 Prioritized memory access explains planning and hippocampal replay. *Nat. Neurosci.* **21**, 1609–1617. (doi:10.1038/s41593-018-0232-z)
108. Kurth-Nelson Z, Behrens T, Wayne G, Miller K, Luettgau L, Dolan R, Liu Y, Schwartenbeck P. 2023 Replay and compositional computation. *Neuron* **111**, 454–469. (doi:10.1016/j.neuron.2022.12.028)

109. Sagiv Y, Akam T, Witten IB, Daw ND. 2025 Between planning and map building: prioritizing replay when future goals are uncertain. *Neuron* **113**, 4278–4292. (doi:10.1016/j.neuron.2025.09.021)
110. van der Meer MA, Bendor D. 2025 Awake replay: off the clock but on the job. *Trends Neurosci.* **48**, 257–267. (doi:10.1016/j.tics.2025.02.006)
111. Schapiro AC, McDevitt EA, Rogers TT, Mednick SC, Norman KA. 2018 Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nat. Commun.* **9**, 3920. (doi:10.1038/s41467-018-06213-1)
112. van de Ven GM, Siegelmann HT, Tolias AS. 2020 Brain-inspired replay for continual learning with artificial neural networks. *Nat. Commun.* **11**, 4069. (doi:10.1038/s41467-020-17866-2)
113. Russek EM, Momennejad I, Botvinick MM, Gershman SJ, Daw ND. 2017 Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Comput. Biol.* **13**, e1005768. (doi:10.1371/journal.pcbi.1005768)
114. Wittkuhn L, Krippner LM, Koch C, Schuck NW. 2025 Replay in the human visual cortex during brief task pauses is linked to implicit learning of successor representations. *Proc. Natl Acad. Sci. USA* **122**, e2507516122. (doi:10.1073/pnas.2507516122)
115. Foster DJ, Wilson MA. 2006 Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440**, 680–683. (doi:10.1038/nature04587)
116. Ambrose RE, Pfeiffer BE, Foster DJ. 2016 Reverse replay of hippocampal place cells is uniquely modulated by changing reward. *Neuron* **91**, 1124–1136. (doi:10.1016/j.neuron.2016.07.047)
117. Singer AC, Frank LM. 2009 Rewarded outcomes enhance reactivation of experience in the hippocampus. *Neuron* **64**, 910–921. (doi:10.1016/j.neuron.2009.11.016)
118. Oudiette D, Paller KA. 2013 Upgrading the sleeping brain with targeted memory reactivation. *Trends Cogn. Sci.* **17**, 142–149. (doi:10.1016/j.tics.2013.01.006)
119. Lewis PA, Bendor D. 2019 How targeted memory reactivation promotes the selective strengthening of memories in sleep. *Curr. Biol.* **29**, R906–R912. (doi:10.1016/j.cub.2019.08.019)
120. Tambini A, Davachi L. 2019 Awake reactivation of prior experiences consolidates memories and biases cognition. *Trends Cogn. Sci.* **23**, 876–890. (doi:10.1016/j.tics.2019.07.008)
121. Bendor D, Wilson MA. 2012 Biasing the content of hippocampal replay during sleep. *Nat. Neurosci.* **15**, 1439–1444. (doi:10.1038/nn.3203)
122. Shearer CM, Rawson AB, Barron HC, O'Reilly JX. 2025 Memory reactivation during rest forms shortcuts in a cognitive map. *Scient. Rep.* **15**, 24724. (doi:10.1038/s41598-025-06742-y)
123. Rawson AB, Nalluru S, O'Reilly JX, Barron HC. 2024 Memory reactivation generates new, adaptive behaviours that reach beyond direct experience. *Scient. Rep.* **14**, 30097. (doi:10.1038/s41598-024-78906-1)
124. Pezzulo G, Zorzi M, Corbetta M. 2021 The secret life of predictive brains: what's spontaneous activity for? *Trends Cogn. Sci.* **25**, 730–743. (doi:10.1016/j.tics.2021.05.007)
125. Bicanski A, Burgess N. 2018 A neural-level model of spatial memory and imagery. *eLife* **7**, e33752. (doi:10.7554/eLife.33752)
126. Becker S, Burgess N. 2001 Modelling spatial recall, mental imagery and neglect. In *Advances in Neural Information Processing Systems 13: Proc. 2000 Conf.* (eds TK Leen, TG Dietterich, V Tresp), pp. 96–102. Cambridge, MA: MIT Press.
127. Byrne P, Becker S, Burgess N. 2007 Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol. Rev.* **114**, 340–375. (doi:10.1037/0033-295X.114.2.340)
128. Madar AD, Jiang A, Dong C, Sheffield MEJ. 2025 Synaptic plasticity rules driving representational shifting in the hippocampus. *Nat. Neurosci.* **28**, 848–860. (doi:10.1038/s41593-025-01894-6)
129. Kesby JP, Eyles DW, McGrath JJ, Scott JG. 2018 Dopamine, psychosis and schizophrenia: the widening gap between basic and clinical neuroscience. *Transl. Psychiatry* **8**, 30. (doi:10.1038/s41398-017-0071-9)
130. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. 2013 The computational anatomy of psychosis. *Front. Psychiatry* **4**, 47. (doi:10.3389/fpsy.2013.00047)
131. Sterzer P *et al.* 2018 The predictive coding account of psychosis. *Biol. Psychiatry* **84**, 634–643. (doi:10.1016/j.biopsych.2018.05.015)
132. Horga G, Abi-Dargham A. 2019 An integrative framework for perceptual disturbances in psychosis. *Nat. Rev. Neurosci.* **20**, 763–778. (doi:10.1038/s41583-019-0234-1)
133. Schlagenhauf F, Huys QJM, Deserno L, Rapp MA, Beck A, Heinze HJ, Dolan R, Heinz A. 2014 Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage* **89**, 171–180. (doi:10.1016/j.neuroimage.2013.11.034)
134. Katthagen T, Fromm S, Wieland L, Schlagenhauf F. 2022 Models of dynamic belief updating in psychosis—a review across different computational approaches. *Front. Psychiatry* **13**, 814111. (doi:10.3389/fpsy.2022.814111)
135. Haddad PM, Correll CU. 2018 The acute efficacy of antipsychotics in schizophrenia: a review of recent meta-analyses. *Ther. Adv. Psychopharmacol.* **8**, 303–318. (doi:10.1177/2045125318781475)
136. Knight S, McCutcheon R, Dwir D, Grace AA, O'Daly O, McGuire P, Modinos G. 2022 Hippocampal circuit dysfunction in psychosis. *Transl. Psychiatry* **12**, 344. (doi:10.1038/s41398-022-02115-5)
137. Mancini V, Delavari F, Eom TY, Zakharenko SS, Schmitt E, Eliez S. 2026 How does altered function of the hippocampus contribute to the development of psychosis? *Biol. Psychiatry* **99**, 113–123. (doi:10.1016/j.biopsych.2025.06.022)
138. Mancini V, Sandini C, Padula MC, Zöller D, Schneider M, Schaer M, Eliez S. 2020 Positive psychotic symptoms are associated with divergent developmental trajectories of hippocampal volume during late adolescence in patients with 22q11DS. *Mol. Psychiatry* **25**, 2844–2859. (doi:10.1038/s41380-019-0443-z)
139. Smeland OB *et al.* 2018 Genetic overlap between schizophrenia and volumes of hippocampus, putamen, and intracranial volume indicates shared molecular genetic mechanisms. *Schizophr. Bull.* **44**, 854–864. (doi:10.1093/schbul/sbx148)
140. Schobel SA *et al.* 2013 Imaging patients with psychosis and a mouse model establishes a spreading pattern of hippocampal dysfunction and implicates glutamate as a driver. *Neuron* **78**, 81–93. (doi:10.1016/j.neuron.2013.02.011)
141. Anticevic A *et al.* 2015 Association of thalamic dysconnectivity and conversion to psychosis in youth and young adults at elevated clinical risk. *JAMA Psychiatry* **72**, 882–891. (doi:10.1001/jamapsychiatry.2015.0566)
142. Harrisberger F *et al.* 2016 Impact of polygenic schizophrenia-related risk and hippocampal volumes on the onset of psychosis. *Transl. Psychiatry* **6**, e868. (doi:10.1038/tp.2016.143)
143. McHugo M, Talati P, Armstrong K, Vandekar SN, Blackford JU, Woodward ND, Heckers S. 2019 Hyperactivity and reduced activation of anterior hippocampus in early psychosis. *Am. J. Psychiatry* **176**, 1030–1038. (doi:10.1176/appi.ajp.2019.19020151)
144. Provenzano FA *et al.* 2020 Hippocampal pathology in clinical high-risk patients and the onset of schizophrenia. *Biol. Psychiatry* **87**, 234–242. (doi:10.1016/j.biopsych.2019.09.022)
145. Mancini V, Saleh MG, Delavari F, Bagautdinova J, Eliez S. 2023 Excitatory/inhibitory imbalance underlies hippocampal atrophy in individuals with 22q11.2 deletion syndrome with psychotic symptoms. *Biol. Psychiatry* **94**, 569–579. (doi:10.1016/j.biopsych.2023.03.021)
146. Lewis DA, Curley AA, Glausier JR, Volk DW. 2012 Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends Neurosci.* **35**, 57–67. (doi:10.1016/j.tics.2011.10.004)

147. Mancini Vet *et al.* 2022 Aberrant developmental patterns of gamma-band response and long-range communication disruption in youths with 22q11.2 deletion syndrome. *Am. J. Psychiatry* **179**, 204–215. (doi:10.1176/appi.ajp.2021.21020190)
148. Thompson M, Weickert CS, Wyatt E, Webster MJ. 2009 Decreased glutamic acid decarboxylase₆₇ mRNA expression in multiple brain areas of patients with schizophrenia and mood disorders. *J. Psychiatr. Res.* **43**, 970–977. (doi:10.1016/j.jpsychi.2009.02.005)
149. Nakazawa K, Zsiros V, Jiang Z, Nakao K, Kolata S, Zhang S, Belforte JE. 2012 GABAergic interneuron origin of schizophrenia pathophysiology. *Neuropharmacology* **62**, 1574–1583. (doi:10.1016/j.neuropharm.2011.01.022)
150. Taylor SF, Tso IF. 2015 GABA abnormalities in schizophrenia: a methodological review of in vivo studies. *Schizophr. Res.* **167**, 84–90. (doi:10.1016/j.schres.2014.10.011)
151. Ramsaran Alet *et al.* 2023 A shift in the mechanisms controlling hippocampal engram formation during brain maturation. *Science* **380**, 543–551. (doi:10.1126/science.ade6530)
152. Lisman JE, Pi HJ, Zhang Y, Otmakhova NA. 2010 A thalamo-hippocampal-ventral tegmental area loop may produce the positive feedback that underlies the psychotic break in schizophrenia. *Biol. Psychiatry* **68**, 17–24. (doi:10.1016/j.biopsych.2010.04.007)
153. Lodge DJ, Grace AA. 2011 Hippocampal dysregulation of dopamine system function and the pathophysiology of schizophrenia. *Trends Pharmacol. Sci.* **32**, 507–513. (doi:10.1016/j.tips.2011.05.001)
154. Dickerson DD, Wolff AR, Bilkey DK. 2010 Abnormal long-range neural synchrony in a maternal immune activation animal model of schizophrenia. *J. Neurosci.* **30**, 12424–12431. (doi:10.1523/JNEUROSCI.3046-10.2010)
155. Sigurdsson T, Stark KL, Karayiorgou M, Gogos JA, Gordon JA. 2010 Impaired hippocampal–prefrontal synchrony in a genetic mouse model of schizophrenia. *Nature* **464**, 763–767. (doi:10.1038/nature08855)
156. Adams RA, Bush D, Zheng F, Meyer SS, Kaplan R, Orfanos S, Marques TR, Howes OD, Burgess N. 2020 Impaired theta phase coupling underlies frontotemporal dysconnectivity in schizophrenia. *Brain* **143**, 1261–1277. (doi:10.1093/brain/awaa035)
157. Speers LJ, Cheyne KR, Cavani E, Hayward T, Schmidt R, Bilkey DK. 2021 Hippocampal sequencing mechanisms are disrupted in a maternal immune activation model of schizophrenia risk. *J. Neurosci.* **41**, 6954–6965. (doi:10.1523/JNEUROSCI.0730-21.2021)
158. Convertino L, Bush D, Zheng F, Adams RA, Burgess N. 2023 Reduced grid-like theta modulation in schizophrenia. *Brain* **146**, 2191–2198. (doi:10.1093/brain/awac416)
159. Lisman JE, Jensen O. 2013 The theta-gamma neural code. *Neuron* **77**, 1002–1016. (doi:10.1016/j.neuron.2013.03.007)
160. Lynn PA, Sponheim SR. 2016 Disturbed theta and gamma coupling as a potential mechanism for visuospatial working memory dysfunction in people with schizophrenia. *Neuropsychiatr. Electrophysiol.* **2**, 7. (doi:10.1186/s40810-016-0022-3)
161. Barr MS, Rajji TK, Zomorodi R, Radhu N, George TP, Blumberger DM, Daskalakis ZJ. 2017 Impaired theta-gamma coupling during working memory performance in schizophrenia. *Schizophr. Res.* **189**, 104–110. (doi:10.1016/j.schres.2017.01.044)
162. Grent-’t-Jong T, Gajwani R, Gross J, Gumley AI, Krishnadas R, Lawrie SM, Schwannauer M, Schultze-Lutter F, Uhlhaas PJ. 2020 Association of magnetoencephalographically measured high-frequency oscillations in visual cortex with circuit dysfunctions in local and large-scale networks during emerging psychosis. *JAMA Psychiatry* **77**, 852–862. (doi:10.1001/jamapsychiatry.2020.0284)
163. Grent-’t-Jong T, Brickwedde M, Metzner C, Uhlhaas PJ. 2023 40-Hz auditory steady-state responses in schizophrenia: toward a mechanistic biomarker for circuit dysfunctions and early detection and diagnosis. *Biol. Psychiatry* **94**, 550–560. (doi:10.1016/j.biopsych.2023.03.026)
164. Uhlhaas PJ, Singer W. 2011 The development of neural synchrony and large-scale cortical networks during adolescence: relevance for the pathophysiology of schizophrenia and neurodevelopmental hypothesis. *Schizophr. Bull.* **37**, 514–523. (doi:10.1093/schbul/sbr034)
165. Mukherjee A, Carvalho F, Eliez S, Caroni P. 2019 Long-lasting rescue of network and cognitive dysfunction in a genetic schizophrenia model. *Cell* **178**, 1387–1402. (doi:10.1016/j.cell.2019.07.023)
166. Strüber M, Sauer JF, Bartos M. 2022 Parvalbumin expressing interneurons control spike-phase coupling of hippocampal cells to theta oscillations. *Scient. Rep.* **12**, 1362. (doi:10.1038/s41598-022-05004-5)
167. Cardin JA, Carlén M, Meletis K, Knoblich U, Zhang F, Deisseroth K, Tsai LH, Moore CI. 2009 Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* **459**, 663–667. (doi:10.1038/nature08002)
168. Suh J, Foster DJ, Davoudi H, Wilson MA, Tonegawa S. 2013 Impaired hippocampal ripple-associated replay in a mouse model of schizophrenia. *Neuron* **80**, 484–493. (doi:10.1016/j.neuron.2013.09.014)
169. Altimus C, Harrold J, Jaaro-Peled H, Sawa A, Foster DJ. 2015 Disordered ripples are a common feature of genetically distinct mouse models relevant to schizophrenia. *Mol. Neuropsychiatry* **1**, 52–59. (doi:10.1159/000380765)
170. Nour MM, Liu Y, Arumham A, Kurth-Nelson Z, Dolan RJ. 2021 Impaired neural replay of inferred relationships in schizophrenia. *Cell* **184**, 4315–4328. (doi:10.1016/j.cell.2021.06.012)
171. Ohki T, Chao ZC, Takei Y, Kato Y, Sunaga M, Suto T, Tagawa M, Fukuda M. 2024 Multivariate sharp-wave ripples in schizophrenia during awake state. *Psychiatry Clin. Neurosci.* **78**, 507–516. (doi:10.1111/pcn.13702)
172. Howes OD, Nour MM. 2016 Dopamine and the aberrant salience hypothesis of schizophrenia. *World Psychiatry* **15**, 3–4. (doi:10.1002/wps.20276)
173. Corlett PR, Fraser KM. 2025 20 years of aberrant salience in psychosis: what have we learned? *Am. J. Psychiatry* **182**, 819–829. (doi:10.1176/appi.ajp.20240556)
174. Murphy KRet *et al.* 2025 A practical guide to transcranial ultrasonic stimulation from the IFCN-endorsed ITRUSST consortium. *Clin. Neurophysiol.* **171**, 192–226. (doi:10.1016/j.clinph.2025.01.004)
175. Xie Z, Dong S, Zhang Y, Yuan Y. 2023 Transcranial ultrasound stimulation at the peak-phase of theta-cycles in the hippocampus improve memory performance. *Neuroimage* **283**, 120423. (doi:10.1016/j.neuroimage.2023.120423)
176. Pan D, D’Ambrogio S, Kingston N, Rascu M, Sankhe P, Luo S, Klein-Flügge MC, Mahmoodi A, Rushworth MFS. 2025 Causal necessity of human hippocampus for structure-based inference in learning. *bioRxiv* 2025.08.19.664920. (doi:10.1101/2025.08.19.664920)
177. Grossman Net *et al.* 2017 Noninvasive deep brain stimulation via temporally interfering electric fields. *Cell* **169**, 1029–1041. (doi:10.1016/j.cell.2017.05.024)
178. Violante IRet *et al.* 2023 Non-invasive temporal interference electrical stimulation of the human hippocampus. *Nat. Neurosci.* **26**, 1994–2004. (doi:10.1038/s41593-023-01456-8)
179. Beanato Eet *et al.* 2024 Noninvasive modulation of the hippocampal-entorhinal complex during spatial navigation in humans. *Sci. Adv.* **10**, eado4103. (doi:10.1126/sciadv.ado4103)