A unifying account of replay as context-driven memory reactivation

Zhenglong Zhou*, Michael J. Kahana†, & Anna C. Schapiro†

Department of Psychology, University of Pennsylvania

*Correspondence: zzhou34@sas.upenn.edu †Equal contribution

Abstract

During rest and sleep, sequential neural activation patterns corresponding to awake experience re-emerge, and this replay has been shown to benefit subsequent behavior and memory. Whereas some studies show that replay directly recapitulates recent experience, others demonstrate that replay systematically deviates from the temporal structure, the statistics, and even the content of recent experience. Given these disparate characteristics, what is the nature and purpose of replay? Here, we offer a theoretical framework in which replay reflects simple context-guided processes that facilitate memory. We suggest that, during awake learning, the brain associates experiences with the contexts in which they are encoded, at encoding rates that vary according to the salience of each experience. During quiescence, replay emerges as the result of a cascade of autonomous bidirectional interactions between contexts and their associated experiences, which in turn facilitates memory consolidation. A computational model instantiating this proposal explains numerous replay phenomena, including findings that existing models fail to account for and observations that have been predominantly construed through the lens of reinforcement learning. Our theory provides a unified, mechanistic framework of how the brain initially encodes and subsequently replays experiences in the service of memory consolidation.

Introduction

Sleep and rest are crucial for learning and memory [1, 2, 3]. A candidate mechanism facilitating these benefits is replay — the offline re-emergence of neural activity associated with awake experience. Over the past several decades, the field of neuroscience has accumulated extensive evidence of replay [4, 5, 6] and increasing evidence of its utility to behavior [7, 8, 9, 10, 11]. It is challenging, however, to characterize the principles and functions of replay, because it exhibits disparate characteristics across states and task contexts that are difficult to synthesize under one framework. Early studies showed that replay preserves the correlational structure and the temporal structure of multi-cell spiking patterns that underlie awake experiences [12, 13, 14]. Most canonically, the firing of sequences of hippocampal place cells corresponding to traversals through an environment re-emerge with the same sequential firing during rest and sleep. Subsequent studies, however, demonstrated that replay deviates from the temporal structure, statistics, and content of recent experience in myriad ways: Replay activates never-experienced novel trajectories [15], over-represents salient experiences [16, 17, 18], unrolls in the reverse order of a behavioral sequence when an animal consumes reward [19, 20, 21, 22], and sometimes exhibits a bias away from recently experienced trajectories [23, 15]. These observations illustrate that replay is not a simple, direct recapitulation of awake experience.

An influential explanation for why replay exhibits distinctive properties, especially in the presence of reward, is that replay is for learning value predictions in the manner of reinforcement learning (RL) models [19, 20, 9, 24, 25]. According to this perspective, adaptive behavior depends on identifying actions that lead to rewarding outcomes, which requires predicting the downstream value of actions (i.e., eventual punishment or reward). This perspective argues that replay reactivates memories to update these predictions. For example, many speculate that reverse replay implements a classic method for updating value predictions, which is to propagate information backward from a reward through experienced trajectories to update upstream actions' value predictions [26]. A recent theory extending this perspective [24] argues that a range of other replay characteristics [19, 27, 28, 18, 17] can be explained by assuming that replay prioritizes updates that will most improve future behavior. The model assumes that replay knows in advance which updates will best improve behavior, defined by a quantity called the expected value of backup (EVB). When replay progresses from updates that improve future behavior the most to the least (i.e., highest to lowest EVB), it produces patterns that match a number of empirically-observed phenomena. However, knowing the behavioral consequence of replay before it is performed is implausible, and the model makes predictions that are inconsistent with empirical data, including the prediction that replay always prioritizes updates relevant to the present goal [23, 29, 15], and that learning reduces the rate of backward more than forward replay [30, 31]. It is thus unclear how strong an explanation the RL perspective provides for the characteristics of replay.

We offer an alternative theoretical account in which replay reflects simple context-guided memory processes. We hypothesize that, during awake learning, the brain sequentially associates experiences with the contexts in which they are encoded, in a manner modulated by the salience of each experience (more rapid association for more salient experiences). During quiescence — both awake rest periods and sleep, replay arises as the result of a cascade of bidirectional interactions between contexts and their associated experiences. The offline brain continues to learn from this replay, updating associations between reactivated memories and contexts. In this account, replay does not compute the utility of memories for learning value predictions nor does it track value predictions. Instead, replay arises naturally from a memory-based mechanism operating bidirectionally between contexts and their associated experiences.

We show that an instantiation of this account — a computational model that builds on established context-based memory encoding and retrieval mechanisms [32, 33] — unifies numerous replay phenomena. These include replay patterns often presumed to involve RL computations and findings that existing models do not account for. First, in the model, the content and structure of replay sequences vary according to states and task contexts in ways that mirror empirical observations [34, 28, 20, 35]. Second, the model captures prominent effects of reward on replay [19, 16, 17] despite not tracking value predictions. Third, in line with

a number of findings [36, 15, 37, 21, 38], replay is not restricted to direct recent experience: The model reactivates non-local and never-experienced novel trajectories. Moreover, the model captures a range of experience-dependent replay characteristics [39, 23, 27, 28, 15, 30, 37, 31]. Finally, replay benefits memory consolidation in ways that align with prior observations and theories [40, 41, 42, 9, 22, 43, 44]. As a whole, our framework provides a general, mechanistic account of how the brain initially encodes and subsequently reactivates experiences in the service of memory consolidation.

Results

A context model of memory replay

Our proposed model builds on memory encoding and retrieval mechanisms established in retrieved-context models of memory, as exemplified in the context-maintenance and retrieval model (CMR; [32, 33]). We refer to the model as CMR-replay (Figure 1b). We begin with an overview of the model architecture, followed by illustrations of awake learning and replay in the model by considering how it encodes and reactivates sequences of items (Figure 1a). We provide a detailed walk-through of operations in the model in Methods.

CMR-replay comprises four elements (Figure 1b): item (f), context (c), item-to-context associations (M^{fc}), and context-to-item associations (M^{cf}). Prior to learning, distinct items (f) are associated with orthogonal context features. As CMR-replay encodes (during awake learning) or reactivates (during replay) a sequence of items, f represents the current item, whereas a drifting context (c) maintains a receding history of context features associated with present and past items (Figure 1c). During both awake learning and replay, CMR-replay associates each item with the current drifting context by updating bidirectional associations between them — item-to-context associations (M^{fc}), which map each item to its associated context, and context-to-item associations (M^{cf}), which map each context to associated items. Prior to learning, these associations are fully orthogonal: M^{fc} maps each item to a distinct context and M^{cf} maps each context feature to a distinct item. The values of the elements of the f and c vectors can be considered to correspond abstractly to the firing rate of a neuron or population of neurons, and the values of M^{cf} and M^{fc} to the strength of synaptic connections between those neurons.

During awake encoding of a sequence of items, for each item, the model retrieves its associated context via M^{fc}. The context layer incorporates the item's associated context and downweights its representation of contexts associated with previous items (Figure 1c). Thus, the context layer maintains a recency weighted sum of contexts associated with past and present items. To perform encoding, CMR-replay updates M^{fc} and M^{cf} to strengthen associations between the current item and context (Figure 1d). The M^{fc} update adds the current context to the context associated with the item. Because the current context contains contexts associated with previous items, through the M^{fc} update, the context associated with the item starts to reflect contexts of prior items. For the same reason, through the M^{cf} update, M^{cf} learns to map contexts associated with previous items to the current item.

Building on prior work [45, 46], CMR-replay embraces the simplifying assumption that the salience of each item influences its rate of encoding (i.e., the learning rates at which the model updates M^{fc} and M^{cf}). In particular, the model updates M^{fc} and M^{cf} at faster rates for salient items, including those that are novel and rewarding (see Methods), than for others. Higher encoding rates allow salient items to form associations with their encoding contexts more rapidly. In CMR-replay, salience modifies encoding rates only for the current item and context: It does not modify encoding rates for the items that lead up to the salient item.

During replay, the model generates a cascade of item and context reactivations by operating bidirectionally on M^{fc} and M^{cf}. This process begins with an initial item reactivation. At the onset of each replay period, the model selects this item from an activity distribution that reflects spontaneous activity during awake rest or sleep (Figure 1e). To capture the near-absence of external sensory information during sleep, we initialized

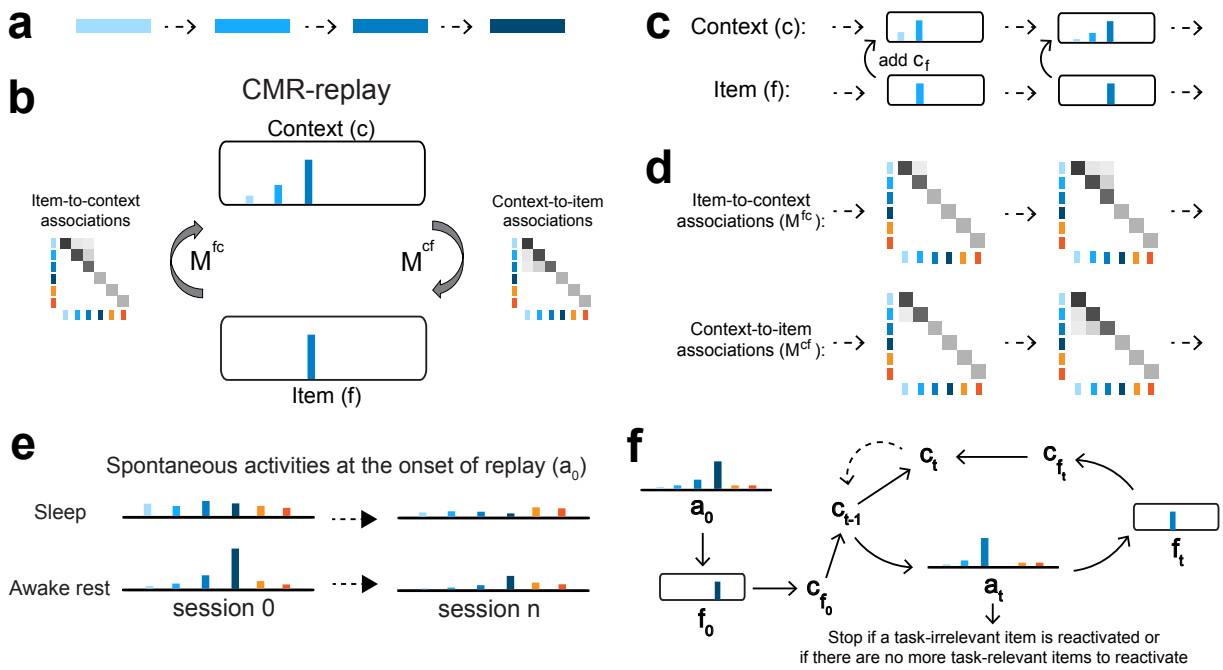

Figure 1: **CMR-replay**. **a**. Consider a task of encoding a sequence consisting of four items, each denoted by a shade of blue. **b**. We propose a model of replay that builds on the context-maintenance and retrieval model (CMR), which we refer to as CMR-replay. The model consists of four components: experience (f), context (c), item-to-context associations (M^{fc}), and context-to-item associations (M^{cf}). At each timestep during awake encoding, f represents the current item and c is a recency-weighted average of associated contexts of past and present items. CMR-replay associates f and c at each timestep, updating M^{fc} and M^{cf} according to a Hebbian learning rule. M^{fc} and M^{cf} respectively support the retrieval of an item's associated context and a context's associated items. During replay, f represents the current reactivated item and c is a drifting context representing a recency-weighted average of associated contexts of past and present reactivated items. Here, too, the model updates M^{fc} and M^{cf} to associate reactivated f and c. The figure illustrates the representations of f, c, M^{fc}, and M^{cf} as the model encodes the third item during learning. Lengths of color bars in f and c represent relative magnitudes of different features. Shades of grey illustrate the weights in M^{fc} and M^{cf}. Orange features represent task-irrelevant items, which do not appear as inputs during awake encoding but compete with task-relevant items for reactivation during replay. **c**. During both awake encoding and replay, context c drifts by incorporating the current item f's associated context c_f and downweighting previous items' associated contexts. The figure illustrates how context drifts during the first time the model encodes the example sequence. **d**. The figure illustrates M^{fc} and M^{cf} updates as the model encodes the third item during the first presentation of the sequence. **e**. Consider the activation of items at the onset of sleep and awake rest across sessions of learning. At replay onset, an initial probability distribution across experiences a_0 varies according to the behavioral state (i.e., awake rest or sleep). Compared to sleep, during awake rest, a_0 is strongly biased toward features associated with external inputs during awake rest. For awake rest, the figure shows an example of a_0 when the model receives a context cue related to the fourth item. Through repeated exposure to the same task sequence across sessions of learning, activities of the four task-related items (i.e., blue items) become suppressed in a_0 relative to task-irrelevant items (i.e., orange items). **f**. Each replay period begins by sampling an experience $f_{t=0}$ according to a_0, where t denotes the current timestep. If $f_{t=0}$ is a task-related item, its associated context $c_{f_{t=0}}$ is reinstated as c_0 to enter a recursive process. During this process, at each timestep $t \geq 1$, c_{t-1} evokes a probability distribution a_t that excludes previously reactivated experiences. Given a_t, the model samples an experience f_t and reinstates f_t's associated context c_{f_t}, which is combined with c_{t-1} to form a new context c_t to guide the ensuing reactivation. The dashed arrow indicates that c_t becomes c_{t-1} for the next time step. At any t, the replay period ends with a probability of 0.1 or if a task-irrelevant item is reactivated.

this distribution with random activity across all items at sleep onset. In simulations of awake rest, in contrast, we present a cue that represents the external sensory context of where the model is "resting." This cue evokes activities that bias the distribution toward this resting location. As a result, awake replay exhibits an initiation bias — a tendency to initiate at the item most strongly associated with the cue.

Once CMR-replay reactivates an initial item, this triggers a sequence of autonomous reactivation (Figure 1f). As in awake encoding, during replay, the model maintains a drifting context — a recency-weighted average of past and present reactivated items' associated contexts. At each timestep, using the current context as a cue, the model evokes a distribution of activities across items via M^{cf}. The model converts these activities into a probability distribution and samples another item without replacement (i.e., by excluding previously reactivated items). In accordance with awake encoding, the model updates context by incorporating the newly reactivated item's associated context, and updates M^{fc} and M^{cf} to associate the reactivated item with the updated context albeit at much slower rates. The updated context then guides item reactivation at the next timestep. At each timestep, a replay event ends with a constant probability or if a task-irrelevant item becomes reactivated.

In the model, replay preserves the temporal contiguity of awake experience, such that each reactivated item tends to be followed by the item that was encoded immediately after or before it (Fig. 6d left). During awake encoding, because of the way context incrementally drifts, the encoding contexts for adjacent items are more similar than for items that are far apart. During replay, when the model retrieves a reactivated item's associated context to guide the next reactivation, it will then favor the reactivation of items that immediately preceded or followed the current item during awake encoding (Fig. 6d left). This behavior is referred to as the model's contiguity bias, which allows replay to generate coherent sequences despite its stochasticity. Following prior work [24], we consider replay events to be replayed sequences (one per replay period) with consecutive segments of length five or greater that preserve the contiguity of the awake sequence.

In memory models that reactivate memories for offline learning [47, 48, 49], a common issue is that the most well-learned items are rehearsed most often, leading to additional strengthening of these items, leading in turn to even more rehearsal, and so on. CMR models can exhibit this same rich-get-richer phenomenon. The solution in prior models has been to incorporate a mechanism that balances rehearsal across items [48, 49, 50]. CMR-replay has such a mechanism as well, which increasingly downweights task-related items at the onset of replay through repetition in the same task. This process downweights items according to the activity of their retrieved contexts in the preceding awake encoding period. As CMR-replay repeatedly strengthens weights for a sequence of items, the probability that replay begins with its constituent items decreases, allowing alternative items that received less exposure to participate in replay (Fig. 1c). The proposal that a suppression mechanism plays a role in replay aligns with models that regulate place cell reactivation via inhibition [51], and empirical observations of increases in hippocampal inhibitory interneuron activity with experience [39]. There are multiple possibilities for how a biological process may implement something like our suppression mechanism [52]. The need across theoretical perspectives for some form of balancing mechanism motivates the strong prediction that such a biological mechanism must be at work.

We simulate awake learning in a number of tasks [19, 34, 23, 28, 15, 21, 9, 17] by exposing the model to sequences of items that correspond to trajectories of spatial locations or visual stimuli as experienced by animals in the experiments. In between sessions of wake learning, we simulate quiescence (both awake rest and sleep) as periods of autonomous reactivation. Our objective is to examine whether CMR-replay can capture qualitative aspects of existing replay phenomena, rather than to provide a quantitative fit to the data. Unlike prior work that finds different best-fitting parameters across simulations [33, 53, 45], CMR-replay employs one set of model parameters across all simulations. In the following sections, we show that the model accounts for a diverse range of empirical phenomena, including context-dependent variations of replay (Fig. 2), effects of reward on replay (Fig. 3), replay patterns that go beyond direct recent experience (Fig. 4), experience-dependent variations of replay (Fig. 5), and ways in which replay facilitates memory (Fig. 6).

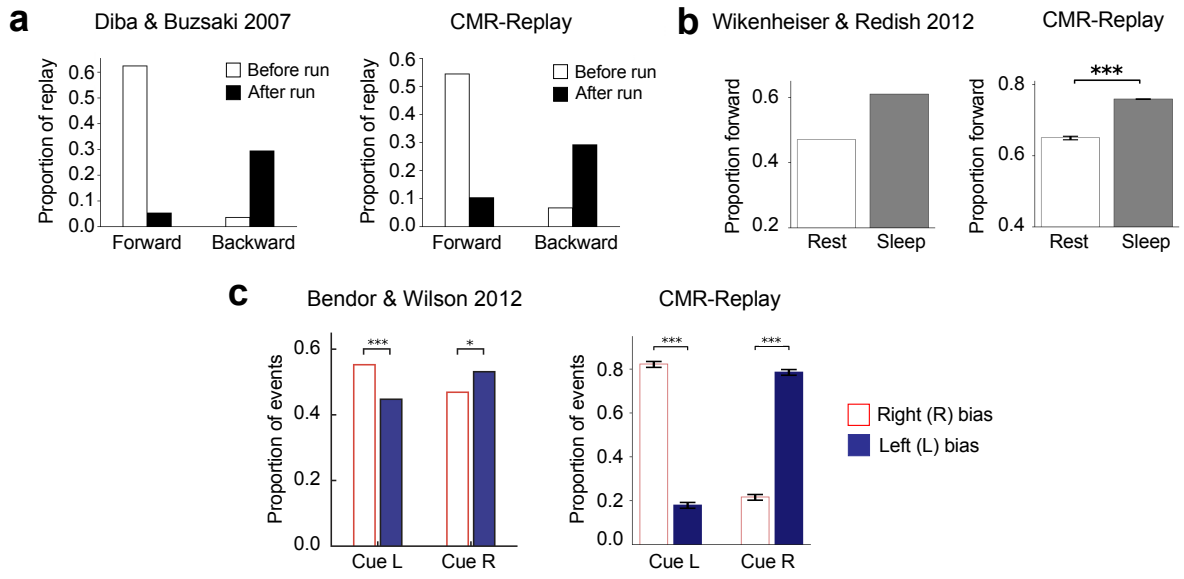

Figure 2: **Context-dependent variations in memory replay**. **a**. As observed in rodents (left), replay in CMR-replay (right) is predominantly forward at the start of a run and backward at the end of a run on a linear track. **b**. Consistent with rodent data (left), in CMR-replay (right), the proportion of forward replay is higher during sleep than during awake rest. **c**. The presence of external cues during sleep biases replay toward their associated memories both in animals (left) and in CMR-replay (right). Error bars represent ±1 standard error of the mean. *p<0.05; **p<0.01; ***p<0.001. Left image in **a** adapted from ref [28], Nature Publishing Group; left image in **b** adapted from ref [35], Wiley; left image in **c** adapted from [34], Nature Publishing Group.

The context-dependency of memory replay

During quiescence, sequential neural firing during sharp-wave ripples (SWRs) recapitulates the temporal pattern of previous waking experience [4]. We distinguish between forward and backward replay, defined as neural activity that either preserves the order of a prior experience (forward replay) or reverses it (backward replay). In animals and humans, the content and directionality of replay systematically vary according to task contexts and behavioral states [34, 28, 21, 35]. For example, animals tend to shift from forward to backward replay from the beginning to the end of a run [28], exhibit more forward replay during sleep [35], and show biased replay of memories associated with external cues during sleep [34]. Some of these observations have led investigators to posit distinct processes underlying forward and backward replay [54, 19, 28, 4, 20, 55, 24, 31], with forward replay supporting planning at choice points [28, 4, 24, 31] and backward replay encoding value expectations from reward outcomes [19, 4, 20]. Here we evaluate whether CMR-replay can account for these differential patterns under one framework, with replay always reflecting associations between items and contexts.

When animals traverse a linear track to collect a reward, forward replay predominates during pre-run rest [28]. In contrast, backward replay predominates during post-run rest, when the animal consumes its reward (see Fig. 2a left; [28]). We simulated this task by presenting CMR-replay with a sequence of items (Fig. 7a), each representing a distinct location. These item representations can be considered to correspond to place cells in rodents, whose activity is typically used to track replay. During post-run rest, we use the final item's encoding context as an external cue for rest replay. For pre-run rest, the first item's encoding context serves as the external cue for rest replay. Because of the external cue, awake rest replay initiates disproportionately at the item most strongly associated with the cue (Fig. 1e), which is consistent with a bias of awake replay to initiate at the resting location [56]. The conjunction of this initiation bias and the model's contiguity bias entails that replay tends to unroll successively forward from the first item in pre-run rest and backward from the final item in post-run rest (Fig. 2a right) as in the data [28]. In contrast to the

EVB model [24], CMR-replay captures the graded nature of this phenomenon (Fig. 2a right): Forward and backward replay appear in both conditions [28]. All of the differences between conditions observed here, and across all simulations in the paper, are highly reliable ($p < 0.001$ based on two-tailed t-tests, with runs of the model as random effects factor). Levels of significance are indicated in figures.

As with prior retrieved context models [32, 33], CMR-replay encodes stronger forward than backward associations. This asymmetry exists because, during the first encoding of a sequence, an item's associated context contributes only to its ensuing items' encoding contexts. Therefore, after encoding, bringing back an item's associated context is more likely to reactivate its ensuing than preceding items, leading to forward asymmetric replay (Fig. 6d left). Absent external cues, sleep replay is less likely to initiate at the final item than rest replay (Fig. 1e), allowing for more forward transitions. This leads to more forward replay during sleep than awake rest (Fig. 2b right), matching empirical observations [4, 57, 35] (Fig. 2b left). In contrast, the EVB model predicts a predominance of reverse replay before behavior stabilizes [24]. Note that the overall proportion of forward replay is higher in the model than these data, but consistent with that found in Diba and Buzsaki (2007).

We next asked whether CMR-replay can simulate Targeted Memory Reactivation (TMR) — the re-presentation of learning-related cues during sleep to encourage reactivation of the associated information. One study employed the TMR paradigm in rodents, associating distinct auditory cues (L and R) with left and right traversal of a linear track [34]. Playing each auditory cue during sleep elicited replay of place cell activity in the cued direction. We simulate these findings by encoding two sequences that share a start item. To simulate TMR, we presented a distinct cue item after each sequence's start item during learning (Fig. 7e), and re-present each cue item (through its associated context) as an external cue in sleep. Matching Bendor and Wilson [34], CMR-replay preferentially replayed each cue's associated sequence (Fig. 2c right).

Effects of reward

At first glance, our proposal may seem at odds with extensive evidence of the influence of reward on replay [19, 20, 30, 21, 9, 22, 16, 17] because CMR-replay neither maintains nor updates value representations during replay. For example, studies suggest that replay over-represents experiences with rewarded or aversive outcomes [30, 16, 58, 18, 17] and awake reverse replay occurs primarily during reward receipt [19, 28, 20]. Reverse replay's unique sensitivity to reward [19, 22] appears to indicate a functional distinction between forward and backward replay, with backward replay specialized for learning value-based predictions [19, 20, 9].

We suggest that salience governs encoding rates, which aligns with evidence that salient stimuli bind more strongly to their context [59, 60, 61, 62]. Building on models that adopt this assumption [45, 46], CMR-replay updates M^{fc} and M^{cf} at higher rates for salient experiences, including those with high valence (reward or punishment) or novelty. In CMR-replay, increasing encoding rates strengthens replay in two distinct ways: Enhancing the M^{cf} encoding rate facilitates the reactivation of an item given features of its encoding context as cue, while enhancing the M^{fc} encoding rate facilitates the faithful retrieval of an item's encoding context. Here, we explore whether these mechanisms allow CMR-replay to account for reward-related phenomena.

After visually exploring a T-maze with one arm containing reward, animals preferentially activated sequences representing the rewarded arm during sleep [17] (Fig. 3a left). We simulated this task by presenting CMR-replay with two sequences, one with a rewarded final item and the other with a neutral final item (Fig. 7d). Due to the influence of encoding rates, replay over-represents the rewarded item compared to the matched neutral item (Fig. 3a right) as in empirical observations [30, 16, 17]. CMR-replay exhibits this property without the assumption that reward-associated items receive more exposure during encoding [63].

Varying the magnitude of reward at the end of a linear track significantly alters the number of backward but not forward replay events [19] (Fig. 3b left). Following Mattar and Daw [24], to disambiguate each

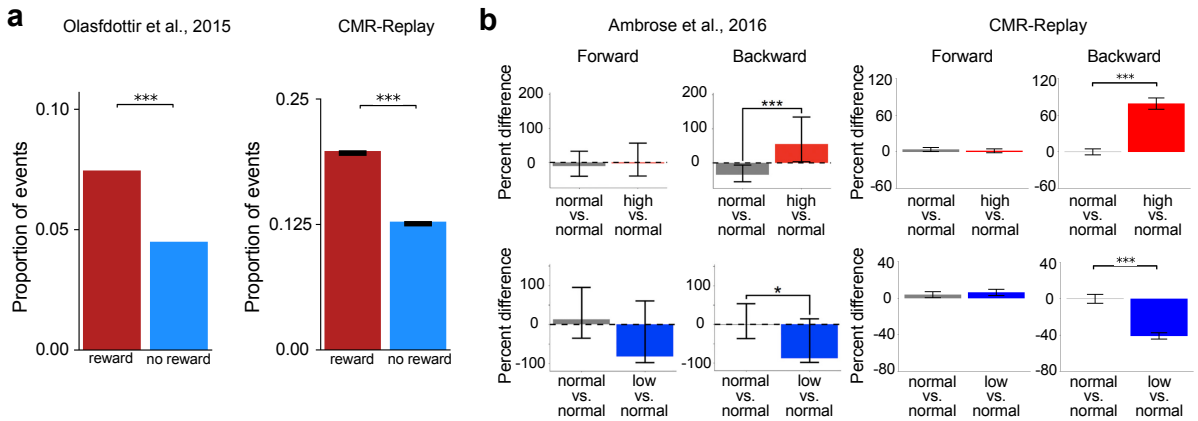

Figure 3: **Reward leads to over-representation in sleep and modulates the rate of backward replay**. **a**. Sleep over-represents experiences associated with reward in animals (left) and in CMR-replay (right). Error bars represent ±1 standard error of the mean. **b**. Varying the magnitude of reward outcome leads to differences in the frequency of backward but not forward replay in animals (left) and CMR-replay (right). In the animal data (left), error bars show 95% confidence intervals. For simulation results (right), error bars show ±1 standard error of the mean. Left image in **a** adapted from ref [17], eLife; images in the left column adapted from ref [19], Elsevier.

location and the direction of a run, we simulated the task with two distinct input sequences, each with a final rewarded item. We manipulated the encoding rate of one rewarded item to be higher (i.e., high reward), lower (i.e., low reward), or identical to that of the reward item in the other sequence (i.e., normal reward). Since the encoding of the rewarded item primarily influences backward replay in post-run rest, we observed differences in the rate of backward but not forward replay between different reward conditions (Fig. 3b right), matching empirical observations [19, 22].

CMR-replay's ability to account for the effects of reward supports our proposal that reward modulates the initial encoding of memories to shape subsequent replay. After reward exerts its influence during encoding, prioritized replay of rewarded memories can occur even if reward-related activity is absent. Consistent with our proposal that replay itself does not require value-based computations, sleep's preferential consolidation of reward memories does not seem to require dopaminergic activity [64], and the coordination between reward responsive neurons and replay-related events is absent in sleep [65]. Our model treats reward as simply the salient feature of an item, generating the prediction that non-reward-related salient items should exhibit similar characteristics.

Replay goes beyond direct recent experience

We next asked whether CMR-replay can account for findings in which animals replay sequences learned outside of their present context. Several studies have established this so-called "remote replay" phenomenon [15, 37]. Here we describe one such experiment and show how CMR-replay provides an account of its findings. In Gupta et al. [15], animals explored both arms of a T-maze during pre-training. During each subsequent recording session, animals traversed only the left or right arm (L- or R- only conditions) or alternated between them (alternation condition). During reward receipt on the just-explored arm, awake rest exhibited remote replay of the opposite, non-local arm (Fig. 4a left: remote replay) across all conditions [15]. This observation challenges models that prioritize items near the resting location [24] and recently active neurons [66, 67, 20, 68] throughout replay. To determine whether CMR-replay can reproduce these results, we presented the model with sequences that overlap for the first few items (representing the central arm of the T-maze; Fig. 7c). During each of two simulated "pre-training" sessions, the model encoded both sequences. We then ran the model through two conditions in an ensuing "experimental" session, where it encoded either

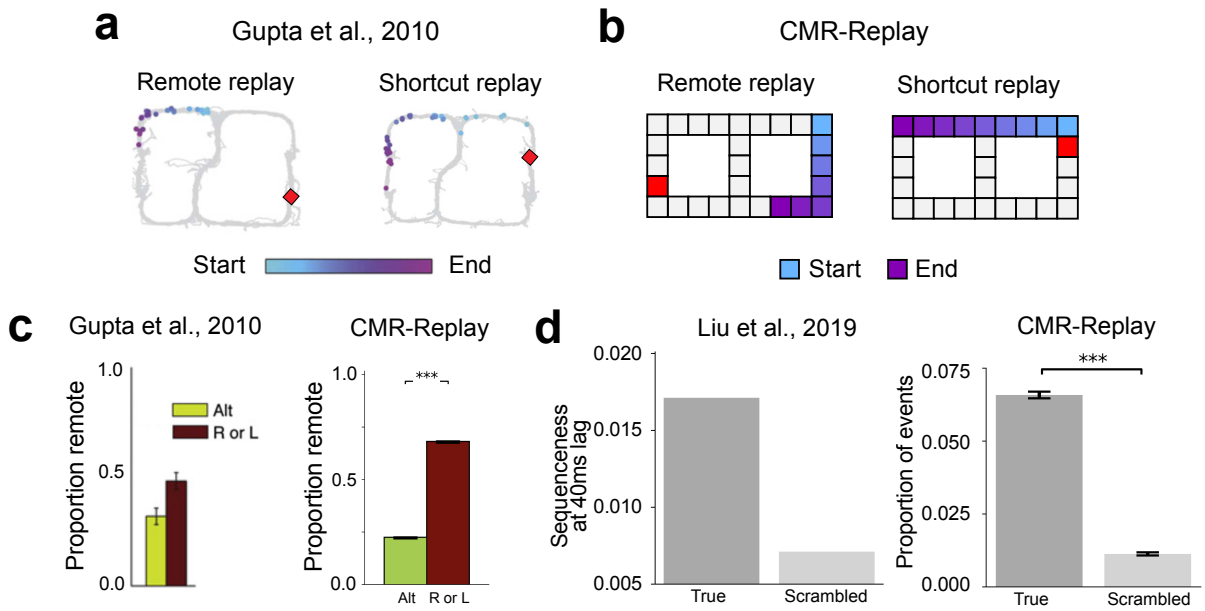

Figure 4: **Replay activates remote experiences and links temporally-separated experiences**. **a**. The two panels show examples of remote and novel (shortcut) replay sequences observed in animals. The colored items indicate the temporal order of the sequences (light blue, early; purple, late). The red item denotes the resting position. **b**. CMR-replay also generates remote and shortcut rest replay, as illustrated in the two predicted sequences of neural firing in the two panels. **c**. Proportion of replay events that contain remote sequences in animals (left) and in CMR-replay (right). Error bars show $\pm$ 1 standard error of the mean in the data and model. **d**. In Liu et al. (2019), participants encoded scrambled versions of two true sequences $X_1X_2X_3X_4$ and $Y_1Y_2Y_3Y_4$: $X_1X_2Y_1Y_2$, $X_2X_3Y_2Y_3$, and $X_3X_4Y_3Y_4$ (Fig. 7g). After learning, human spontaneous neural activity showed stronger evidence of sequential reactivation of the true sequences (left). CMR-replay encoded scrambled sequences as in the experiment. Consistent with empirical observation, subsequent replay in CMR-replay over-represents the true sequences (right). Error bars show $\pm$ 1 standard error of the mean in the model. Images in **a** adapted from ref [15], Elsevier; left image in **c** adapted from ref [15], Elsevier; left image in **d** adapted from ref [21], Elsevier.

only one (L or R -only conditions) or both sequences (alternation condition). After encoding the sequences, we simulated reward receipt by presenting CMR-replay with the encoding context of a rewarded item as an external context cue. As in Gupta et al. [15], CMR-replay is able to generate remote replay of the non-local sequence (Fig. 4b, left; Fig. 4c, right). When CMR-reactivates a non-local item by chance, replay context dramatically shifts by incorporating the non-local item's associated context, thereby triggering a cascade of non-local item reactivation to generate remote replay. Due to its suppression mechanism, CMR-replay is able to capture the higher prevalence of remote replay in L and R -only conditions (Fig. 4c, right), which we will unpack in a subsequent section. The occurrence of remote replay does not require the suppression mechanism in CMR-replay, as the model generates remote replay in the alternation condition where suppression is matched across local and non-local items.

We next examined whether replay in CMR-replay can link temporally-separated experiences to form novel sequences that go beyond direct experience [36, 63, 15, 69, 21]. Gupta et al. [15] showed the occurrence of novel, shortcut replay sequences. These shortcut sequences cut directly across the choice point and link segments of the two arms of a T-maze during rest (Fig. 4a right: shortcut replay), even though animals never directly experienced such trajectories [15]. In our simulation of the study [15], CMR-replay also generates novel rest replay that links segments of the two sequences (Fig. 4b, right): The reactivation of the juncture of the two sequences (the top middle item of Fig. 7c) brings back context common to the two sequences, allowing replay to stitch together segments of the two sequences. In line with Gupta et al. [15], shortcut replay appeared at very low rates in CMR-replay (the alternation condition: mean propor-

tion of replay events that contain shortcut sequence = 0.0046; L or R conditions: mean proportion = 0.0062).

Liu et al. [21] showed that replay in humans reorganizes temporally-separated wake inputs. In their first experiment, participants encoded sequences that scrambled pairwise transitions of two true sequences $X_1X_2X_3X_4$ and $Y_1Y_2Y_3Y_4$: $X_1X_2Y_1Y_2$, $X_2X_3Y_2Y_3$, and $X_3X_4Y_3Y_4$. To highlight transitions from the true sequences, the time lag between those transitions (e.g., X_2X_3) was shorter than others (e.g., X_3Y_2) during presentation. Analyses revealed preferential replay of the true as opposed to the scrambled sequences [21] (Fig. 4d, left). We simulated the experiment by presenting CMR-replay with sequences of the same structure (Fig. 7g), where context drifted more for longer interstimulus intervals. After learning, the model performed replay in the absence of external context cues. The quantification of replay in Liu et al. [21], which reflects statistical evidence of replay decoding based on magnetoencephalography (MEG) data, differs from our measure, where we have direct access to replay without measurement noise. However, qualitatively matching Liu et al. [21], CMR-replay preferentially replays true sequences relative to scrambled sequences (Fig. 4d, right).

The influence of experience

Task exposure influences replay, with replay appearing less frequently in familiar as compared with novel environments [27, 70, 71, 68]. Task repetition similarly reduces replay [28, 31]. After gaining experience along multiple trajectories, animals and humans can exhibit enhanced replay of non-recently explored trajectories [23, 29, 15, 38]. Overall, these findings demonstrate a negative relationship between the degree and recency of experience and the frequency of replay. This pattern challenges models in which experience monotonically enhances the reactivation of items [66, 67, 63, 20, 24, 68].

In CMR-replay, experience shapes replay in two opposing ways. First, repetition strengthens M^{fc} and M^{cf}, allowing replay to better preserve the temporal structure of waking inputs. Second, by enhancing M^{fc}, repetition increases the activity of contexts associated with items. Since CMR-replay suppresses the activity of items at the onset of replay as a function of their activity during learning, repetition increases the downweighting of the activity of task items, reducing their probability of reactivation. Such a suppression mechanism may be adaptive in allowing replay to benefit not just the most recently- and/or strongly- encoded items [72].

The next set of simulations illustrates CMR-replay's account of experience-dependent changes in replay [23, 28, 15, 30, 31]. We first examined how replay changes through repeated encoding of the same inputs following our linear track simulation illustrated in Fig. 7a. Here, CMR-replay encodes the same sequence across learning sessions, with awake rest after each session. Initially, experience increases the prevalence of replay (Fig. 5a: left). As repetition enhances the suppression of task-related items at the onset of replay, replay frequency subsequently decreases in CMR-replay (Fig. 5a: left). Through experience, the average length of replay increases (Fig. 5a: middle), suggesting that repetition strengthens sequence memory in the model. In contrast to the EVB model [24], which predicts a differential drop in the rate of backward relative to forward replay, the proportion of replay events that are backward does not decrease (Fig. 5a right) in CMR-replay. This result highlights that, unlike the EVB model, CMR-replay does not employ distinct variables to drive forward versus backward replay.

In an experiment where animals learned the same task across eight behavioral sessions, Shin et al. [31] observed a similar pattern of results. As shown in Figure 5b, animals exhibited lower rates of replay but longer replay sequences in later sessions (left, middle). As in our CMR-replay simulations, as the rates of forward and backward replay both decrease, the proportion of forward relative to backward replay events remains relatively stable across sessions (right). Furthermore, consistent with reduced reactivation of task-related units in CMR-replay, the study observed decreased reactivation of task-related place cells through experience. In contrast, item reactivation increases monotonically through repetition in other models [63, 24]. Shin et al. [31] performed Bayesian decoding to statistically quantify evidence of replay, whereas our analyses

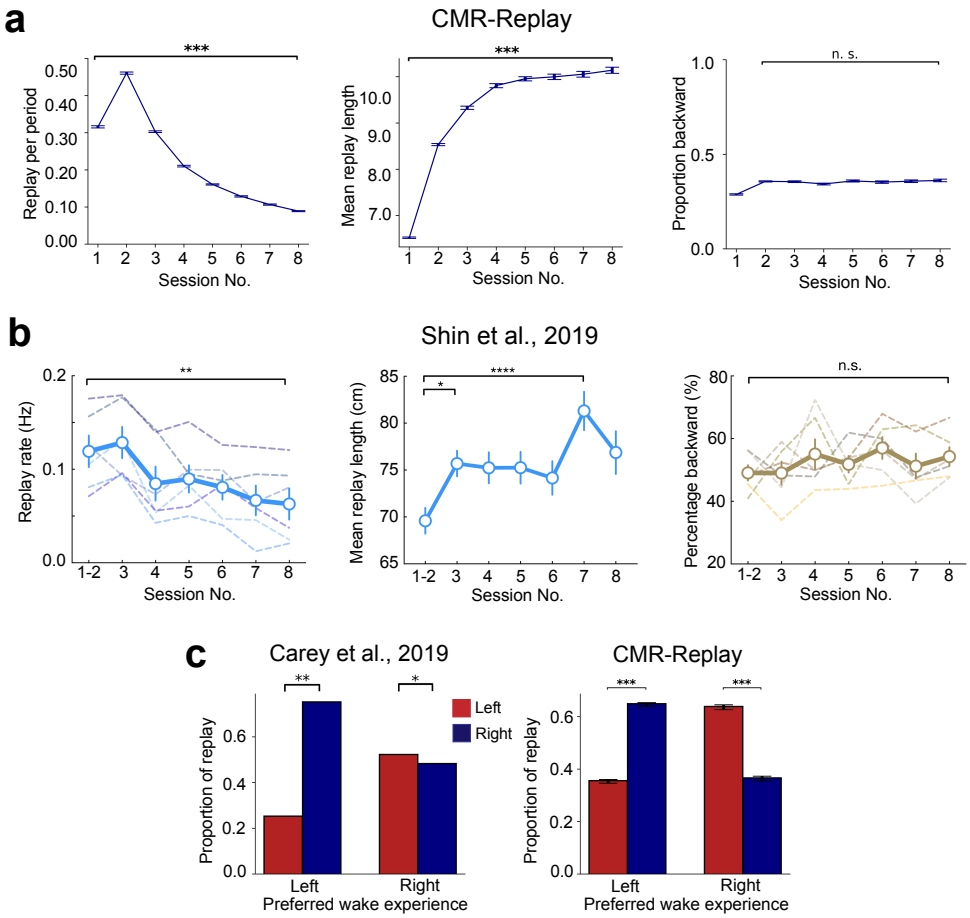

Figure 5: **Variations in replay as a function of experience. a**. In CMR-replay, through repeated exposure to the same task, the frequency of replay events decreases (left), the average length of replay events increases (middle), and the proportion of replay events that are backward remains stable (after a slight initial uptick; right). **b**. With repeated experience in the same task, animals exhibit lower rates of replay (left) and longer replay sequences (middle), while the proportion of replay events that are backward stays relatively stable (right). **c**. In a T-maze task, where animals display a preference for traversing a particular arm of the maze, replay more frequently reflects the opposite arm [23] (left). CMR-replay preferentially replays the right arm after exposure to the left arm and vice versa (right). Error bars show $\pm$ 1 SEM in all panels. Images in **b** adapted from ref [31], Elsevier; left image in **c** adapted from [23], Nature Publishing Group.

directly compare segments of a behavioral sequence with replay sequences. Despite differences between these measures, the patterns of results in the data and in the model match qualitatively. Several other studies using varied experimental procedures have reported similar effects of repeated experience on replay, including a reduction in the prevalence of replay [27, 28], an increase in replay length [39], and no reduction in the proportion of replay events that are backward [30].

In CMR-replay, the activity of retrieved contexts associated with items in a learning session modulates the level of item suppression during ensuing quiescence. As a result, items that get more exposure in a session may receive more suppression than others at the onset of replay, facilitating the reactivation of their competitors. In our simulation of [15] (Fig. 7c), in the L and R -only conditions, since the sequence presented during learning receives more suppression, remote replay is more prevalent than in the alternation condition, where both sequences appear during learning (Fig. 4b). In the L or R -only conditions, when CMR-replay performs post-learning replay in the absence of external context cues, replay over-represents the alternative

sequence (Fig. 5c), which aligns with the observation that replay exhibits a bias away from the arm of a T-maze that animals preferred during behavior [23]. This property is also consonant with recent findings that replay preferentially activates non-recent trajectories [29].

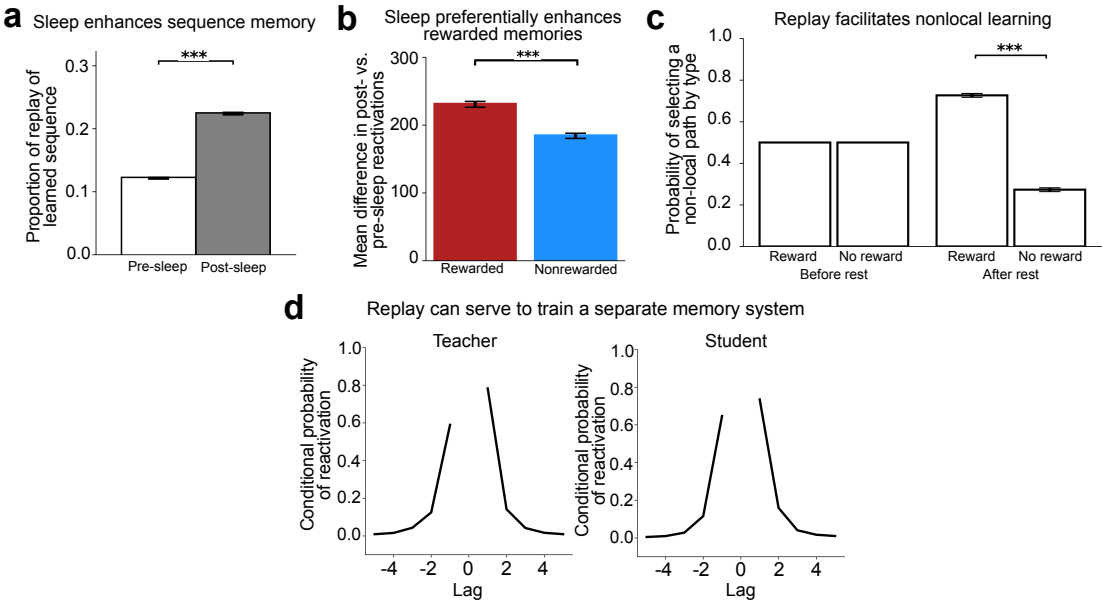

Figure 6: **Learning from replay**. **a**. Sleep increased the likelihood of reactivating the learned sequence in the correct temporal order in CMR-replay, as seen in an increase in the proportion of replay for learned sequences post-sleep. **b**. Sleep leads to greater reactivation of rewarded than non-rewarded experiences, indicating that sleep preferentially strengthens rewarded memories in CMR-replay. **c**. In the simulation of Liu et al. [9], CMR-replay encoded six sequences, each of which transitioned from one of three start items to one of two end items. After receiving a reward outcome for the end item of a sequence, we simulated a period of rest. After but not before rest, CMR-replay exhibited a preference for non-local sequences that led to the rewarded item. This preference emerged through rest despite the fact that the model never observed reward in conjunction with those non-local sequences, suggesting that rest replay facilitates non-local learning in the model. **d**. We trained a "teacher" CMR-replay model on a sequence of items. After encoding the sequence, the teacher generated replay sequences during sleep. We then trained a separate blank-slate "student" CMR-replay model exclusively on the teacher's sleep replay sequences. To assess knowledge of the original sequence, we collected sleep replay sequences from both models, and assessed the probability that each model reactivates the item at position i + lag of the sequence immediately following the reactivation of the i-th item of the sequence, conditioned on the availability of the i-th item for reactivation. Both models demonstrated a tendency to reactivate the item that immediately follows or precedes the just-reactivated item on the original sequence. This result suggests that the student acquired knowledge of the temporal structure of original sequence by encoding only the teacher's replay sequences. Error bars show $\pm$ 1 SEM.

The function of replay

Many have proposed adaptive functions for replay, including for memory consolidation [41, 1, 73], retrieval [41, 11, 74], credit assignment [19, 20, 25], and planning [75, 76, 77]. Growing causal evidence suggests that replay benefits memory: TMR enhances memory [78], and disrupting SWRs impairs memory [8, 10, 11]. Replay facilitates offline learning in our model by updating M^{fc} and M^{cf} according to the internally-reactivated items and contexts during replay. In the following set of simulations, we characterize ways in which replay facilitates memory in the model.

One of the most robust benefits of sleep is on sequence memory, often studied with motor sequence paradigms [42]. To simulate the impacts of sleep replay on sequence memory, we presented CMR-replay with a five-item sequence and examined whether sleep enhanced memory of the sequence. Before and after sleep, we assessed the proportion of replay sequences that matched the input sequence. The assessment occurred in "test"

periods, where learning rates were set to zero and external cues were absent. In post-sleep test, CMR-replay generated a higher proportion of sequences matching the correct sequence than in pre-sleep test (Fig. 6a), indicating that sleep enhances sequence memory in the model.

Replay preferentially enhances rewarded memories [22], and sleep preferentially consolidates salient experiences [43, 44]. In our simulation of a T-maze with reward in one of the two arms [17], we also included pre- and post- sleep test periods to assess how sleep in CMR-replay shapes rewarded versus non-rewarded memory. Through sleep, CMR-replay exhibited a greater increase in its reactivation of the rewarded item compared to a matched neutral item (Fig. 6b), suggesting that sleep preferentially enhances memory associations for rewarded items in CMR-replay.

A recent study [9] presented evidence that replay facilitates non-local value learning. Human participants first learned about the structure of six sequences, each of which begins with one of three start items and terminates with one of two end items. Then, for two sequences that share a start state, participants learned that only one of them leads to reward. After a period of rest during which replay was measured with MEG, among the other four sequences (i.e., the non-local sequences), participants exhibited a behavioral preference for sequences that terminate in the end item associated with reward, despite no direct recent experience with reward in that sequence. The authors suggested that, in accordance with the RL perspective, replay propagates value to associated items, allowing participants to select non-local sequences associated with reward without direct experience. In our simulation of this paradigm, CMR-replay encoded six sequences of the same structure (Fig. 7b), with increased encoding rates to simulate reward receipt, as in the simulations above. To simulate awake rest after reward receipt, we presented the encoding context of the rewarded item as an external cue. Before and after rest, we examined the model's preference among the four non-local sequences, by assessing how much the model activated each non-local start item's two ensuing items given the start item's associated context as a cue. After but not before rest, CMR-replay preferentially activated the item that leads to the rewarded end item (Fig. 6c). This is because the presence of the rewarded item as an external cue evokes the reactivation of its associated non-local items. In CMR-replay, this preference emerged without value updates, suggesting that replay can facilitate nonlocal learning by re-organizing memory associations.

There has been much interest in the memory literature in the possibility that hippocampal replay serves to train neocortical systems to represent recent memories [79, 40, 66, 41, 52, 73, 50, 80]. We explored whether replay in CMR-replay can serve to transfer one model's knowledge to another. After a "teacher" CMR-replay encodes a sequence, we collected its sleep replay sequences to train a blank-slate "student" CMR-replay at replay's learning rates. Through this process, the student inherited the contiguity bias of the teacher (Fig. 6d), suggesting it acquired knowledge of the structure of the teacher's training sequence. This simulation provides a proof of concept that replay in CMR-replay can serve to facilitate memory transfer across systems, in addition to promoting local learning.

Discussion

What is the nature and function of neural replay? We suggest a simple memory-focused framework that explains a wide array of replay phenomena. First, the brain associates experiences with their encoding contexts at rates that vary according to the salience of each experience. Then, in quiescence, the brain replays by spontaneously reactivating a memory and retrieving its associated context to guide subsequent reactivation. Learning continues to occur during these endogenously generated reactivation events. A model embodying these ideas — CMR-replay — explains many qualitative empirical characteristics of replay and its impacts, including patterns previously interpreted as features of reinforcement learning computations, and observations that prior models do not explain.

First, CMR-replay demonstrates basic properties of replay that other models exhibit (or could easily accommodate) [63, 55, 69, 24, 81], including replay's recapitulation of the temporal pattern of past experience during rest and sleep [28, 35], bias toward memories associated with external cues [34], and ability to stitch together temporally-separated experiences to form novel sequences [15, 21]. Second, CMR-replay captures findings that have been interpreted as evidence that replay serves value-based reinforcement learning, including over-representation of memories associated with reward [17], reverse replay upon reward receipt [28, 20], and the unique sensitivity of reverse replay to reward magnitude [19]. Third, CMR-replay accounts for observations that are not naturally accounted for by prior models, including a stable proportion of backward replay through learning [31], reduced item reactivation and sequential replay with experience [31, 27], increased prevalence of forward replay in sleep [35], enhanced replay outside of the current context [15], and a tendency for replay to cover non-behaviorally-preferred experiences [23]. Finally, replay facilitates memory in CMR-replay in ways that align with empirical findings [42, 9, 22, 44, 43]. These include the improvement of sequence memory, preferential strengthening of rewarded memories, facilitation of non-local learning, and endogenous training of a separate memory system in the absence of external inputs.

The EVB model and CMR-replay offer different types of explanation for why replay exhibits its disparate characteristics: The EVB model provides an explicitly normative explanation, whereas CMR-replay offers a mechanistic account. The different levels of explanation raises the possibility that CMR-replay could be considered a mechanistic implementation of EVB. Indeed, there are several shared properties between the models [28, 17, 20, 19]. However, as discussed above, CMR-replay captures observations that appear inconsistent with the EVB model, including the prevalence of non-local replay, the decoupling of replay from behavioral preference, and similar proportions of forward and backward replay over time. In addition to these existing observations, the two models make distinct predictions that can be tested experimentally. For example, the EVB model predicts that reward only modulates the rate of replay when it informs potential change in behavior, whereas CMR-replay considers reward as a salient feature that enhances replay by facilitating associations between item and context. Thus, in scenarios where animals cannot choose between competing actions (e.g., when there are only deterministic paths), CMR-replay but not the EVB model predicts that reward manipulation would lead to changes in the rate of replay. CMR-replay also predicts that salient non-rewarding events should lead to similar patterns of replay as rewarded events. While not explicitly normative, CMR-replay does offer an explanation of the goal of replay, which is to improve memory through offline local learning and systems interactions. CMR-replay posits that replay may facilitate building a stable, unbiased understanding of the environment useful for many different possible future tasks, some of which may be difficult for an animal to predict and therefore optimize in advance.

An ongoing debate concerns to what extent awake replay reflects a process of planning that simulates future scenarios in support of immediate decision-making [77, 75, 76], versus to what extent it serves to store, update, and maintain memory without directly guiding behavior [7, 10, 82]. Evidence supporting the planning hypothesis comes from studies that demonstrate enhanced replay of upcoming behavioral trajectories [77, 83]. However, in tasks that track representations of multiple temporally- and spatially- separated experiences, animals exhibit replay that appears to be decoupled from their behavioral preference [23, 29, 15]. Our model aligns more with the memory perspective, as it is able to capture existing findings without positing that replay serves to optimize behavioral outcome. However, replay of this kind could at times be read out and used by downstream decision-making systems. For example, recent work argues that the dynamics of the retrieval processes in this class of models could support adaptive choice in sequential decision tasks [84]. Overall, our framework argues that replay characteristics are primarily driven by memory principles, and that replay serves to strengthen and reorganize memories, which benefits subsequent — but not necessarily immediate — behavior [29, 82].

Many memory consolidation theories are aligned with CMR-replay in suggesting that replay actively strengthens and re-organizes memories [85, 40, 66, 86, 1, 73, 50]. Contextual binding theory [87], however, takes a different approach, suggesting that residual encoding-related activity elicits merely epiphenomenal replay as

context drifts during quiescence. Our theory echoes this perspective in characterizing replay as an outcome of context-guided processes. However, we diverge from the perspective in suggesting that the emergent replay does significantly benefit memory by strengthening learned associations between items and contexts. Our proposal aligns with a recent TMR study showing that the recapitulation of items' associated contexts during sleep drives changes in memory in humans [88]. Our model also captures observations of enhanced replay of infrequent and remote experiences, which are in tension with the perspective that replay is primarily guided by recent activity [85].

Our model has mainly considered replay occurring during sharp wave ripples [14]. During active behavior in rodents, ordered place cell sequences also activate during the theta oscillation (theta sequences) [89]. Similar to ripple-based replay, theta sequences manifest in both forward- and reverse-order [90], initiate at the animal's location, extend further into upcoming locations through experience [91, 92, 93, 94], cluster around behaviorally-relevant items [95], and have been proposed to correspond to cued memory retrieval [96]. These parallels lead us to speculate that the context-driven mechanisms we have laid out for findings of replay mainly during sharp wave ripples may also be relevant in understanding theta sequences, though future work will be needed to extend the model into this domain.

Another important area for future work is to investigate the mapping between the components of CMR-replay and neural circuitry. Our model employs a series of bidirectional operations between context and item representations to generate replay. These operations might be implemented within the recurrent connections of CA3 in the case of temporally-compressed sharp wave ripple replay. It is possible that these interactions could also play out across the "big loop" of the hippocampus [69] or within cortical circuits [97, 98, 99, 100], which could correspond to slower forms of replay [9, 101]. In quiescence, we posit that the hippocampus can serve as a "teacher" that endogenously samples memory sequences to help establish these associations in neocortical areas, with local context-item loops within the teacher and student areas. This process may be most likely to take place during NREM sleep, when ripples, spindles, and slow oscillations may coordinate replay between the hippocampus and neocortical areas [1]. In alignment with the observation that disrupting entorhinal cortex input to the hippocampus affects only awake replay whereas manipulating hippocampal subfield CA3 activity affects both awake and sleep replay [102], in CMR-replay, the principal distinction between awake rest and sleep is whether external inputs bias replay. There are likely other variables, such as task engagement [103], that modulate the influence of external inputs on replay. Finally, CMR-replay employs a suppression mechanism that facilitates the reactivation of items that received less exposure. This approach builds on prior models that use balancing mechanisms to regulate offline learning [48, 49, 50], but it is not yet clear how the brain may implement such a mechanism.

There exists a range of computational models that simulate replay at different levels of biological detail [104, 54, 105, 24, 106, 107, 108], account for different features of replay [63, 55, 69, 76, 24, 106, 81], and posit distinct functions for replay [109, 110, 76, 24, 106, 73, 111, 50]. CMR-replay provides a high-level description of a mechanism that accounts for replay phenomena without simulating realistic spiking, synaptic, or membrane potential mechanisms as in alternative models [104, 54, 112, 113]. Aspects of the model, such as its lack of regulation of the amount of positive weight change that can accumulate from prolonged replay, are biologically implausible and leave opportunities for future work spanning levels of biological detail. However, a diversity of models at different levels of description is useful for the field as we engage with empirical data at different levels (e.g. spikes vs. behavior). Our theory follows a lineage of memory-focused replay models, demonstrating the power of this perspective in accounting for data that have been assumed to require optimization of value-based predictions. As CMR-replay builds on existing theories of memory retrieval, our account is in line with recent proposals that reactivation and retrieval may have similar underlying mechanisms and utility for behavior [114]. In sum, our theory unifies a wealth of phenomena, offering an integrative and mechanistic framework characterizing how the brain initially encodes and subsequently replays memories to facilitate behavior.

Methods

Representation and Initialization

CMR-replay comprises four components as in previous retrieved-context models [32, 33]: item (f), context (c), item-to-context associations (M^{fc}), and context-to-item associations (M^{cf}). During both awake encoding and replay, f represents the current item (i.e., an external input presented during awake learning or a reactivated item during replay). c represents a recency-weighted sum of contexts associated with past and present items. During both awake encoding and replay, CMR-replay associates each item with its current encoding context by updating two sets of weights M^{fc} and M^{cf}. M^{fc} represents item-to-context associations that support the retrieval of an item's associated context. M^{cf} represents context-to-item associations that enable the retrieval of a context's associated items.

CMR-replay employs a one-hot representation of f (i.e., a localist item representation): Each item is represented by a vector of length n in which only the unit representing the item is on (i.e., has an activity of 1) and all other units are off. As illustrated in Fig. 1, in addition to task-related items shown as inputs during learning, we include task-irrelevant items that do not appear during learning but compete with encoded items for reactivation during replay. We use n_{task}, $n_{non-task}$, and n to respectively denote the number of task-related items, the number of task-irrelevant items, and the total number of items (i.e., the sum of n_{task} and $n_{non-task}$). To allow for sufficient competition between task-related and task-irrelevant items, we set $n_{non-task}$ to be roughly one half of n_{task} (i.e., rounded up when n_{task} is odd). We note that the particular ratio of $n_{non-task}$ to n_{task} is not critical to the pattern of results in our simulations.

In each simulation, M^{fc} and M^{cf} are initialized as identity matrices of rank n, which are scaled respectively by 1.0 and 0.7. These scaling factors were chosen to qualitatively match the empirically observed proportions of forward and backward replay in different conditions (though the forward/backward asymmetry is always observed in the model). Our initialization of these two matrices as identity matrices differs from the initialization strategy in prior work [45, 32, 53, 33], where M^{fc} and M^{cf} are initialized to reflect pre-experimental similarity among items. Given such initializations, prior to learning, M^{fc} maps distinct items onto orthogonal context features, and M^{cf} maps each context feature to a different item. Before the model encodes each input sequence, c is reset as a zero vector of length n. Resetting contexts in between sequence presentations demarcates boundaries between discrete events as in prior work [115].

Awake encoding

Context drift

During awake encoding, the model encodes a sequence of distinct items by associating them with a drifting context. At each timestep t, CMR-replay retrieves the current item f_t's associated context c_{f_t} via item-to-context matrix M^{fc}_{t-1} (i.e., the M^{fc} matrix after the previous timestep) according to:

$$c_{f_t} = \frac{M^{fc}_{t-1} f_t}{\parallel M^{fc}_{t-1} f_t \parallel} \tag{1}$$

Given c_{f_t}, the context layer incorporates c_{f_t} and downweights previous items' associated contexts according to:

$$c_t = \rho c_{t-1} + \beta c_{f_t} \tag{2}$$

where c_{t-1} represents the context before f_t is presented. ρ and β determine the relative contribution of c_{t-1} and c_{f_t} to c_t. To ensure that c_t has a unit length, ρ is computed according to:

$$\rho = \sqrt{1 + \beta^2[(c_{t-1}c_{f_t})^2 - 1]} - \beta(c_{t-1}c_{f_t}) \tag{3}$$

Therefore, the context layer is a drifting, recency-weighted average of contexts associated with items presented up to timestep t. Operations that drive context drift in our simulations, including those specified by Eqs. 1, 2, and 3, are identical to those in prior work [45, 33, 53]. In all simulations, β is 0.75 (similar to drift rates for temporal context features reported in Polyn et al. [33]), except when distractors cause context drift in the simulation of Liu et al [21].

Updating M^{fc} and M^{cf}

Each time the context drifts, CMR-replay updates M^{fc} and M^{cf} to strengthen associations between the current item (f_t) and context (c_t). The model updates M^{fc} and M^{cf} using a standard Hebbian learning rule according to

$$\triangle M^{fc} = \gamma_{fc} c_t f_t^T \tag{4}$$

$$\triangle M^{cf} = \gamma_{cf} f_t c_t^T \tag{5}$$

in which γ_{fc} and γ_{cf} control the rates at which M^{fc} and M^{cf} are updated.

Varying awake encoding rates according to salience

CMR-replay assumes that salience modulates the magnitude of γ_{fc} and γ_{cf}. Building on prior work [45, 46], the model assumes that γ_{fc} and γ_{cf} for awake encoding are higher for salient items, including those that are rewarding (i.e., directly associated with a reward) or novel (i.e., received less exposure), than for other items. Salience modulates γ_{fc} and γ_{cf} only for the current item and context: It does not modify γ_{fc} and γ_{cf} for items that precede the salient item. Higher encoding rates allow salient items to associate with their encoding contexts more quickly than other items. This shapes subsequent replay in two distinct ways. First, with higher M^{cf} encoding rate for a salient item, the model will more strongly activate the item given its encoding context as a cue after awake encoding. Second, with higher M^{fc} encoding rate, the model will be better able to faithfully retrieve the item's encoding context after awake encoding. For all simulations, the base learning rates γ_{fc} and γ_{cf} are 1.0. For rewarded items, learning rates vary according to the magnitude of reward: Learning rates $\gamma_{fc_{\text{low}}}$ and $\gamma_{cf_{\text{low}}}$ are 1.0 for experiences associated with low reward, $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$ are 1.5 for those with standard reward, and $\gamma_{fc_{\text{high}}}$ and $\gamma_{cf_{\text{high}}}$ are 2.0 for those with high reward. These values are chosen to align with prior work [45, 46], in which these scaling factors are 1.0 for items that evoke no arousal [45] and greater than 1.0 for those assumed to evoke emotional arousal [45, 46]. When an input becomes less novel as it is repeated across sessions, its learning rate in the present session is $\frac{\gamma}{i}$, where i is the index of the current session and γ is its initial learning rate.

Replay

After each session of awake encoding, CMR-replay autonomously reactivates items during a number of replay periods. For simplicity, we assumed that the number of replay periods is fixed, rather than determined by task-related variables.

Initial item reactivation

At the onset (i.e., $t = 0$) of each replay period, CMR-replay selects an item from a probability distribution a_0, which represents spontaneous activities across items during awake rest or sleep (Fig. 1e). To simulate the relative lack of external sensory input in sleep, we fill a_0 with random activities across all items. By contrast, for awake rest simulations, we make available an external context cue $c_{external}$ (e.g., the task context of the animal's current location) representing where the model is "resting," which evokes additional activities that bias a_0 toward the item most strongly associated with the context cue. Concretely, in a_0, the activity of i-th unit is:

$$[a_0]_i = \frac{[{a_0}^{random} + \lambda {a_0}^{evoked}]_i}{\sum_{j=1}^{n} [{a_0}^{random} + \lambda {a_0}^{evoked}]_j} \tag{6}$$

where ${a_0}^{random}$ represents internal activity that we simulated as random noise, ${a_0}^{evoked}$ is activity that $c_{external}$ evokes according to Eq. 7, and n is the total number of items in a simulation. ${a_0}^{random}$ is a vector of size n whose elements are independently and uniformly drawn from the interval $[0, 0.001]$. CMR-replay samples an item f_0 from a_0 for the initial item reactivation.

Subsequent reactivations

Given f_0, the model reinstates its associated context as c_0 according to Eq. 1 (Figure 1f). This allows the model to engage in a series of autonomous reactivations (Figure 1f).

At each timestep $t \geq 1$, CMR-replay reactivates another item without replacement (i.e., by excluding items reactivated at previous timesteps). In particular, the model uses the previous context c_{t-1} as a context cue to evoke a probability distribution a_t that excludes items reactivated prior to t. Let U_t denote the set of items that have not yet been reactivated. The probability of each item in U_t is:

$$[a_t^{evoked}]_i = \frac{\exp([M^{cf} c_{cue}]_i / T_t)}{\sum_{\forall f_j \in U_t} \exp([M^{cf} c_{cue}]_j / T_t)} \tag{7}$$

where T_t is a temperature parameter that scales the relative difference of activities in ${a_t}^{evoked}$. T_0 is 0.1 and T_t is 0.14 for all $t \geq 1$. For $t \geq 1$, $c_{cue} = c_{t-1}$ and $a_t = {a_t}^{evoked}$. In contrast, at $t = 0$, $c_{cue} = c_{external}$ and a_0 is a combination of ${a_0}^{random}$ and ${a_0}^{evoked}$. Based on Eq. 7, all U_t items have non-zero activity in a_t. From a_t, the model samples an item f_t.

Given f_t, the model performs three operations that parallel the operations performed during awake encoding. First, it reinstates f_t's associated context c_{f_t} via M^{fc} according to Eq. 1. Then, c_{f_t} induces a drift in context according to Eq. 2, forming a new context c_t, which will guide the reactivation at the next timestep. Finally, the model strengthens the association between f_t and the current context c_t by updating M^{fc} and M^{cf} according to Eq. 4 and Eq. 5. Compared to awake encoding, the model performs these updates at a slower learning rate γ_{replay} of 0.001 during replay. This slower learning rate allows replay to preserve and strengthen memories despite the noisy nature of replay sequences.

At each timestep, the replay period ends with a probability of 0.1 or if a task-unrelated experience becomes reactivated. Following prior work [24], we consider replayed sequences (one per replay period) with consecutive segments of length five or greater that preserve the contiguity of wake inputs as replay events.

An experience-dependent suppression mechanism

CMR-replay employs a mechanism that suppresses the activity of items in a_0 according to the magnitude of context activity in the preceding awake encoding period. This mechanism differs from related mechanisms

in prior work [53, 33], which scale the degree of competition among items during recall. For each item f presented in a wake learning session, its activity in a_0 is multiplied by:

$$\omega = exp(-||c_f||) \tag{8}$$

where $||c_f||$ is the Euclidean norm of the item's retrieved context vector c_f in the recent wake learning session given by:

$$||c_f|| = \sqrt{\sum_{i=1}^{m} {c_f}_i^2} \tag{9}$$

where m is the number of units in c_f. For items not shown in the wake session, C is 0.0 and thus ω is 1.0.

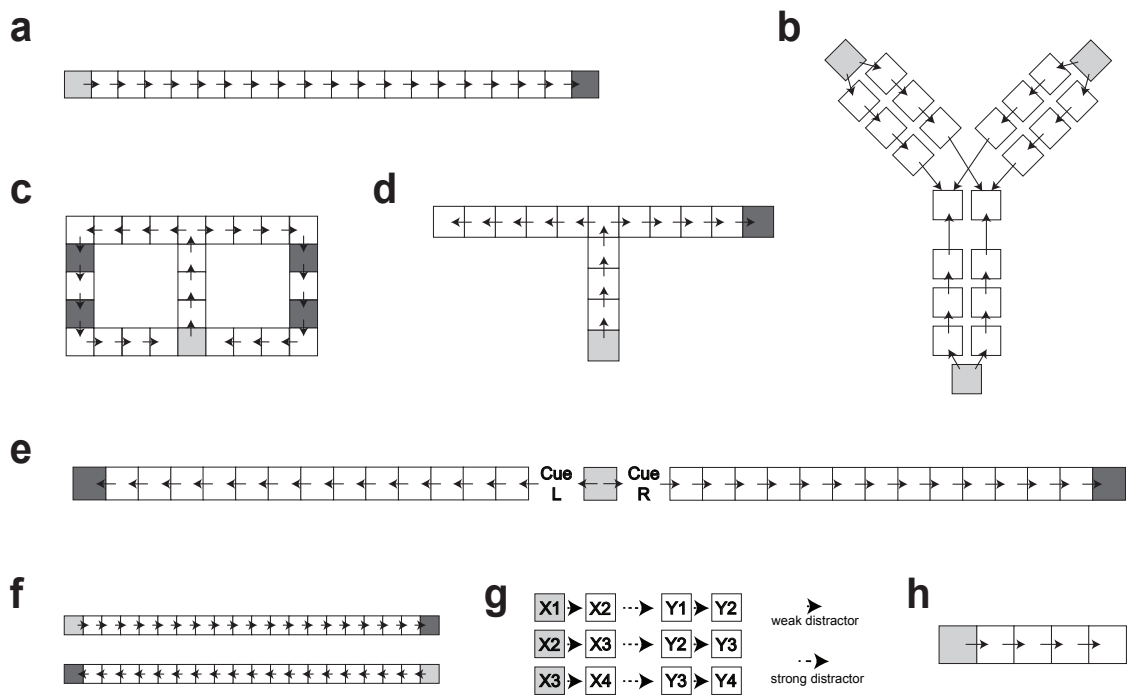

Figure 7: **Task simulations**. Each enclosed box corresponds to a unique item. Each arrow represents a valid transition between two items. Each dashed arrow represents a distractor that causes a drift in context between two items. Task sequences initiate at light grey boxes. Dark grey boxes represent salient items in each task. For tasks with multiple valid sequences, the order in which sequences are presented is randomized. **a**. Simulation of a linear track. **b**. Simulation of the task in Liu et al. [9]. **c**. Simulation of a two-choice T-maze. **d**. Simulation of a T-maze. **e**. Simulation of the task in Bendor and Wilson [34]. **f**. Simulation of a linear track task with distinct directions of travel. **g**. Simulation of input sequences in Liu et al. [21]. **h**. Simulation of a fix-item sequence.

Task Simulations

During awake learning, CMR-replay encodes sequences of items, representing spatial trajectories or other stimulus sequences (Fig. 7). After awake encoding, the model participates in a number of awake rest or sleep replay periods. In each simulation, for each condition, we ran 100 instantiations of the model with the same initialization. Variability in replay sequences across models arises from the stochastic nature of the

replay process. Due to the variability in replay sequences, different models develop distinct M^{fc} and M^{cf} as they learn from replay. Unlike prior work that identified the best-fitting parameters for each simulation [33, 53, 45], CMR-replay employs the same set of model parameters across simulations with varying input structures.

In the simulation that examines context-dependent forward and backward replay through experience (Figs. 2a and 5a), across a total of 8 sessions of awake learning, CMR-replay encodes an input sequence shown in Fig. 7a, which simulates a linear track run with no ambiguity in the direction of inputs. In this simulation, learning rates for the rewarded item are $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$. After each wake learning session, we simulate 500 awake rest replay periods at the end of a run followed by another 500 periods at the start of a run. For rest at the end of a run, $c_{external}$ is the context associated with the final item in the sequence. For rest at the end of a run, $c_{external}$ is the context associated with the start item.

In the simulation that contrasts forward and backward replay between rest and sleep (Fig. 2b), the model encodes the input sequence shown in Fig. 7a for a single session. After encoding, each model either participates in 1000 awake rest or sleep replay periods, with 100 models in each condition (i.e., awake rest or sleep).

In the simulation of TMR (Fig. 2c), each model encodes two sequences shown in Fig. 7e in a randomized order in a session of wake learning. During input presentation, for each sequence, a separate cue item (cue L or cue R) is presented immediately after the start item. The models encode the goal item at rates $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$. After wake learning, each model engages in 500 sleep replay periods. In each replay period, the context associated with cue L or cue R is randomly presented as $c_{external}$.

In the simulation that contrasts replay of rewarded versus non-rewarded experiences (Fig. 3a and 6b), each model encodes two sequences shown in Fig. 7d in a randomized order in a session of wake learning. The models encode the goal item at rates $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$. After wake learning, each model engages in an extended phase with 5000 sleep replay periods. To quantify changes in memory through sleep, in each model, we additionally simulated 5000 replay periods before and after extended sleep with no learning (i.e., M^{fc} and M^{cf} are not updated) and no $c_{external}$.

In the simulation of forward and backward replay with different levels of reward (Fig. 3b), the model encodes two sequences (Fig. 7f) in a randomized order in a single session. The inclusion of two disjoint sequences follows the approach in [24], which simulates different directions of travel to distinguish place cells with directional preference for replay decoding in animal studies. The simulation consists of three conditions: normal vs. normal reward, low vs. normal reward, and high vs. normal reward. In the normal vs. normal condition, each model encodes goal locations in both sequences at rates $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$. In the low vs. normal condition, each model encodes the goal location at rates $\gamma_{fc_{\text{low}}}$ and $\gamma_{cf_{\text{low}}}$ for one sequence and at rates $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$ for the other. Finally, in the high vs. normal condition, each model encodes the goal location at rates $\gamma_{fc_{\text{high}}}$ and $\gamma_{cf_{\text{high}}}$ for one sequence and at rates $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$ for the other. After encoding a sequence, we simulate 500 awake rest replay periods at the end of a run followed by another 500 at the start of a run.

In the simulation of remote replay, shortcut replay, and the over-representation of non-behaviorally-preferred experiences in replay (Figs. 4b, 4c, and 5c), each model encodes two sequences (Fig. 7c) in a randomized order in a total of three sessions. Learning rates for each goal location are $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$. In these simulations, we treat the first two sessions as the period in which an animal is pre-trained extensively on the task. After wake learning in the third session, for the results shown in Figs. 4a and b, each model engages in 500 awake rest replay periods at each of the four goal locations in a randomized order. For the results shown in Figs. 5c, to simulate replay away from the task environment, each model engages in 500 replay periods with no external context cue.

In the simulation of Liu et al. [21] (Fig. 4d), each model encodes three sequences (Fig. 7g) shown in a randomized order. These three sequences $X_1X_2Y_1Y_2$, $X_2X_3Y_2Y_3$, and $X_3X_4Y_3Y_4$ are scrambled versions of pairwise transitions from true sequences $X_1X_2X_3X_4$ and $Y_1Y_2Y_3Y_4$. A distractor item, which is a distinct item that does not participate in replay, induces context drift between successive items. The item induces context drift at a β of 0.3 for transitions that exist in the true sequences and at a β of 0.99 for transitions that do not exist in the true sequences (simulating the longer interstimulus interval between these transitions in the experiment).

In the simulation that examines sequence memory through sleep (Fig. 6a), each model encodes a five-item sequence (Fig. 7h) in a session. After wake learning, each model participates in an extended period of sleep with 5000 replay periods. In each model, we additionally simulated 5000 replay periods before and after extended sleep with no learning (i.e., M^{fc} and M^{cf} are not updated) and no $c_{external}$.

In the simulation that examines replay's role in non-local learning (Fig. 6c), each model encodes six sequences (Fig. 7b) in a randomized order in a session. Sequences in this simulation consist of three start states and two end states. Each start state has a unique sequence that connects it to each of the two end states. The model encodes the final item in the final sequence at rates $\gamma_{fc_{\text{high}}}$ and $\gamma_{cf_{\text{high}}}$ and encodes all other items at base learning rates γ_{fc} and γ_{cf}. After the encoding of all six sequences, each model participates in 5000 awake rest replay periods with the final item's associated context as $c_{external}$.

In the simulation of teacher and student CMR-replay (Fig. 6d), each teacher model encodes a sequence (Fig. 7a) in a session. Teacher models encode the goal location at learning rates $\gamma_{fc_{\text{normal}}}$ and $\gamma_{cf_{\text{normal}}}$. After wake learning, we simulate an extended period of sleep with 5000 replay periods in each model. We then present each teacher model's 5000 replayed sequences as inputs to train a different blank-slate student model with learning rates γ_{replay}.

References

1. Klinzing, J. G., Niethard, N., and Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature neuroscience*, 22(10):1598–1610.

2. Walker, M. P. and Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron*, 44(1):121–133.

3. Wamsley, E. J. (2019). Memory consolidation during waking rest. *Trends in cognitive sciences*, 23(3):171–173.

4. Foster, D. J. (2017). Replay comes of age. *Annual review of neuroscience*, 40:581–602.

5. Liu, Y., Nour, M. M., Schuck, N. W., Behrens, T. E., and Dolan, R. J. (2022). Decoding cognition from spontaneous neural activity. *Nature Reviews Neuroscience*, 23(4):204–214.

6. Ólafsdóttir, H. F., Bush, D., and Barry, C. (2018). The role of hippocampal replay in memory and planning. *Current Biology*, 28(1):R37–R50.

7. Dupret, D., O'Neill, J., Pleydell-Bouverie, B., and Csicsvari, J. (2010). The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nature neuroscience*, 13(8):995–1002.

8. Ego-Stengel, V. and Wilson, M. A. (2010). Disruption of ripple-associated hippocampal activity during rest impairs spatial learning in the rat. *Hippocampus*, 20(1):1–10.

9. Liu, Y., Mattar, M. G., Behrens, T. E., Daw, N. D., and Dolan, R. J. (2021). Experience replay is associated with efficient nonlocal learning. *Science*, 372(6544):eabf1357.

10. Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., and Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature neuroscience*, 12(10):1222–1223.

11. Jadhav, S. P., Kemere, C., German, P. W., and Frank, L. M. (2012). Awake hippocampal sharp-wave ripples support spatial memory. *Science*, 336(6087):1454–1458.

12. Wilson, M. A. and McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, 265(5172):676–679.

13. Skaggs, W. E. and McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*, 271(5257):1870–1873.

14. Nádasdy, Z., Hirase, H., Czurkó, A., Csicsvari, J., and Buzsáki, G. (1999). Replay and time compression of recurring spike sequences in the hippocampus. *Journal of Neuroscience*, 19(21):9497–9507.

15. Gupta, A. S., van der Meer, M. A., Touretzky, D. S., and Redish, A. D. (2010). Hippocampal replay is not a simple function of experience. *Neuron*, 65(5):695–705.

16. Singer, A. C. and Frank, L. M. (2009). Rewarded outcomes enhance reactivation of experience in the hippocampus. *Neuron*, 64(6):910–921.

17. Ólafsdóttir, H. F., Barry, C., Saleem, A. B., Hassabis, D., and Spiers, H. J. (2015). Hippocampal place cells construct reward related sequences through unexplored space. *Elife*, 4:e06063.

18. Wu, C.-T., Haggerty, D., Kemere, C., and Ji, D. (2017). Hippocampal awake replay in fear memory retrieval. *Nature neuroscience*, 20(4):571–580.

19. Ambrose, R. E., Pfeiffer, B. E., and Foster, D. J. (2016). Reverse replay of hippocampal place cells is uniquely modulated by changing reward. *Neuron*, 91(5):1124–1136.

20. Foster, D. J. and Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084):680–683.

21. Liu, Y., Dolan, R. J., Kurth-Nelson, Z., and Behrens, T. E. (2019). Human replay spontaneously reorganizes experience. *Cell*, 178(3):640–652.

22. Michon, F., Sun, J.-J., Kim, C. Y., Ciliberti, D., and Kloosterman, F. (2019). Post-learning hippocampal replay selectively reinforces spatial memory for highly rewarded locations. *Current Biology*, 29(9):1436–1444.

23. Carey, A. A., Tanaka, Y., and van Der Meer, M. A. (2019). Reward revaluation biases hippocampal replay content away from the preferred outcome. *Nature neuroscience*, 22(9):1450–1459.

24. Mattar, M. G. and Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature neuroscience*, 21(11):1609–1617.

25. Momennejad, I., Otto, A. R., Daw, N. D., and Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. *Elife*, 7:e32548.

26. Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning*, 3:9–44.

27. Cheng, S. and Frank, L. M. (2008). New experiences enhance coordinated neural activity in the hippocampus. *Neuron*, 57(2):303–313.

28. Diba, K. and Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature neuroscience*, 10(10):1241–1242.

29. Gillespie, A. K., Maya, D. A. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., Roumis, D. K., Eden, U. T., and Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron*, 109(19):3149–3163.

30. Igata, H., Ikegaya, Y., and Sasaki, T. (2021). Prioritized experience replays on a hippocampal predictive map for learning. *Proceedings of the National Academy of Sciences*, 118(1):e2011266118.

31. Shin, J. D., Tang, W., and Jadhav, S. P. (2019). Dynamics of awake hippocampal-prefrontal replay for spatial learning and memory-guided decision making. *Neuron*, 104(6):1110–1125.

32. Howard, M. W. and Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of mathematical psychology*, 46(3):269–299.

33. Polyn, S. M., Norman, K. A., and Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological review*, 116(1):129.

34. Bendor, D. and Wilson, M. A. (2012). Biasing the content of hippocampal replay during sleep. *Nature neuroscience*, 15(10):1439–1444.

35. Wikenheiser, A. M. and Redish, A. D. (2013). The balance of forward and backward hippocampal sequences shifts across behavioral states. *Hippocampus*, 23(1):22–29.

36. Barron, H. C., Reeve, H. M., Koolschijn, R. S., Perestenko, P. V., Shpektor, A., Nili, H., Rothaermel, R., Campo-Urriza, N., O'Reilly, J. X., Bannerman, D. M., Behrens, T. E., and Dupret, D. (2020). Neuronal computation underlying inferential reasoning in humans and mice. *Cell*, 183(1):228–243.

37. Karlsson, M. P. and Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature neuroscience*, 12(7):913–918.

38. Wimmer, G. E., Liu, Y., McNamee, D., and Dolan, R. J. (2023). Distinct replay signatures for prospective decision-making and memory preservation. *Proceedings of the National Academy of Sciences*, 120(6).

39. Berners-Lee, A., Feng, T., Silva, D., Wu, X., Ambrose, E. R., Pfeiffer, B. E., and Foster, D. J. (2022). Hippocampal replays appear after a single experience and incorporate greater detail with more experience. *Neuron*, 110(11):1829–1842.

40. Born, J. and Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological research*, 76(2):192–203.

41. Carr, M. F., Jadhav, S. P., and Frank, L. M. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature neuroscience*, 14(2):147–153.

42. King, B. R., Hoedlmoser, K., Hirschauer, F., Dolfen, N., and Albouy, G. (2017). Sleeping on the motor engram: the multifaceted nature of sleep-related motor memory consolidation. *Neuroscience & Biobehavioral Reviews*, 80:1–22.

43. Payne, J. D. and Kensinger, E. A. (2010). Sleep's role in the consolidation of emotional episodic memories. *Current Directions in Psychological Science*, 19(5):290–295.

44. Payne, J. D., Stickgold, R., Swanberg, K., and Kensinger, E. A. (2008). Sleep preferentially enhances memory for emotional components of scenes. *Psychological science*, 19(8):781–788.

45. Cohen, R. T. and Kahana, M. J. (2022). A memory-based theory of emotional disorders. *Psychological Review*, 129(4):742.

46. Talmi, D., Lohnas, L. J., and Daw, N. D. (2019). A retrieved context model of the emotional modulation of memory. *Psychological Review*, 126(4):455.

47. Ans, B. and Rousset, S. (2000). Neural networks with a self-refreshing memory: knowledge transfer in sequential learning tasks without catastrophic forgetting. *Connection science*, 12(1):1–19.

48. Meeter, M. (2003). Control of consolidation in neural networks: Avoiding runaway effects. *Connection Science*, 15(1):45–61.

49. Norman, K. A., Newman, E. L., and Perotte, A. J. (2005). Methods for reducing interference in the complementary learning systems model: oscillating inhibition and autonomous memory rehearsal. *Neural Networks*, 18(9):1212–1228.

50. Singh, D., Norman, K. A., , and Schapiro, A. C. (2022). A model of autonomous interactions between hippocampus and neocortex driving sleep-dependent memory consolidation. *Proceedings of the National Academy of Sciences*, 119(44):e2123432119.

51. Malerba, P., Krishnan, G. P., Fellous, J.-M., and Bazhenov, M. (2016). Hippocampal CA1 ripples as inhibitory transients. *PLoS computational biology*, 12(4):e1004880.

52. Hasselmo, M. E., Wyble, B. P., , and Wallenstein, G. V. (1996). Encoding and retrieval of episodic memories: role of cholinergic and gabaergic modulation in the hippocampus. *Hippocampus*, 6(6):693–708.

53. Lohnas, L. J., Polyn, S. M., and Kahana, M. J. (2015). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological review*, 122(2):337.

54. Koene, R. A. and Hasselmo, M. E. (2008). Reversed and forward buffering of behavioral spike sequences enables retrospective and prospective retrieval in hippocampal regions ca3 and ca1. *Neural Networks*, 21(2-3):276–288.

55. Khamassi, M. and Girard, B. (2020). Modeling awake hippocampal reactivations with model-based bidirectional search. *Biological Cybernetics*, 114(2):231–248.

56. Davidson, T. J., Kloosterman, F., and Wilson, M. A. (2009). Hippocampal replay of extended experience. *Neuron*, 63(4):497–507.

57. Lee, A. K. and Wilson, M. A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron*, 36(6):1183–1194.

58. Sterpenich, V., van Schie, M. K. M., Catsiyannis, M., Ramyead, A., Perrig, S., Yang, H.-D., Ville, D. V. D., and Schwartz, S. (2021). Reward biases spontaneous neural reactivation during sleep. *Nature communications*, 12(1):4162.

59. Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of experimental psychology: General*, 134(2):258.

60. MacKay, D. G., Hadley, C. B., and Schwartz, J. H. (2005). Relations between emotion, illusory word perception, and orthographic repetition blindness: Tests of binding theory. *The Quarterly Journal of Experimental Psychology Section A*, 58(8):1514–1533.

61. MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., and Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & cognition*, 32(3):474–488.

62. Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, 2(1):33–52.

63. Diekmann, N. and Cheng, S. (2023). A model of hippocampal replay driven by experience and environmental structure facilitates spatial learning. *bioRxiv*.

64. Asfestani, M. A., Brechtmann, V., Santiago, J., Peter, A., Born, J., and Feld, G. B. (2020). Consolidation of reward memory during sleep does not require dopaminergic activation. *Journal of Cognitive Neuroscience*, 32(9):1688–1703.

65. Gomperts, S. N., Kloosterman, F., and Wilson, M. A. (2015). VTA neurons coordinate with the hippocampal reactivation of spatial experience. *Elife*, 4:e05360.

66. Buzsáki, G. (1989). Two-stage model of memory trace formation: a role for "noisy" brain states. *Neuroscience*, 31(3):551–570.

67. Csicsvari, J., O'Neill, J., Allen, K., and Senior, T. (2007). Place-selective firing contributes to the reverse-order reactivation of CA1 pyramidal cells during sharp waves in open-field exploration. *European Journal of Neuroscience*, 26(3):704–716.

68. O'Neill, J., Senior, T. J., Allen, K., Huxter, J. R., and Csicsvari, J. (2008). Reactivation of experience-dependent cell assembly patterns in the hippocampus. *Nature neuroscience*, 11(2):209–215.

69. Kumaran, D. and McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychological review*, 119(3):573.

70. Giri, B., Miyawaki, H., Mizuseki, K., Cheng, S., and Diba, K. (2019). Hippocampal reactivation extends for several hours following novel experience. *Journal of Neuroscience*, 39(5):866–875.

71. Gorriz, M. H., Takigawa, M., and Bendor, D. (2023). The role of experience in prioritizing hippocampal replay. nature communications. *Nature Communications*, 14(1):8157.

72. Schapiro, A. C., McDevitt, E. A., Rogers, T. T., Mednick, S. C., and Norman, K. A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nature communications*, 9(1):1–11.

73. McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102(3):419.

74. Wimmer, G. E., Liu, Y., Vehar, N., Behrens, T. E., and Dolan, R. J. (2020). Episodic memory retrieval success is associated with rapid replay of episode content. *Nature neuroscience*, 23(8):1025–1033.

75. Mattar, M. G. and Lengyel, M. (2022). Planning in the brain. *Neuron*.

76. Jensen, K. T., Hennequin, G., and Mattar, M. G. (2023). A recurrent network model of planning explains hippocampal replay and human behavior. *bioRxiv*.

77. Pfeiffer, B. E. and Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79.

78. Oudiette, D. and Paller, K. A. (2013). Upgrading the sleeping brain with targeted memory reactivation. *Trends in cognitive sciences*, 17(3):142–149.

79. Alvarez, P. and Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: a simple network model. *Proceedings of the national academy of sciences*, 91(15):7041–7045.

80. Sun, W., Advani, M., Spruston, N., Saxe, A., and Fitzgerald, J. E. (2021). Organizing memories for generalization in complementary learning systems. *BioRxiv*.

81. McNamee, D. C., Stachenfeld, K. L., Botvinick, M. M., and Gershman, S. J. (2021). Flexible modulation of sequence generation in the entorhinal–hippocampal system. *Nature neuroscience*, 24(6):851–862.

82. Joo, H. R. and Frank, L. M. (2018). The hippocampal sharp wave–ripple in memory retrieval for immediate use and consolidation. *Nature Reviews Neuroscience*, 19(12):744–757.

83. Xu, H., Baracskay, P., O'Neill, J., and Csicsvari, J. (2019). Assembly responses of hippocampal CA1 place cells predict learned behavior in goal-directed spatial tasks on the radial eight-arm maze. *Neuron*, 101(1):119–132.

84. Zhou, C. Y., Talmi, D., Daw, N. D., and Mattar, M. G. (2023). Episodic retrieval for model-based evaluation in sequential decision tasks. *PsyArXiv*.

85. Antony, J. W. and Schapiro, A. C. (2019). Active and effective replay: systems consolidation reconsidered again. *Nature Reviews Neuroscience*, 20(8):506–507.

86. Cowan, E. T., Schapiro, A. C., Dunsmoor, J. E., and Murty, V. P. (2021). Memory consolidation as an adaptive process. *Psychonomic Bulletin & Review*, 28:1–15.

87. Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., and Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nature Reviews Neuroscience*, 20(6):364–375.

88. Schechtman, E., Heilberg, J., and Paller, K. A. (2023). Memory consolidation during sleep involves context reinstatement in humans. *Cell reports*, 42(4).

89. Johnson, A. and Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 27(45):12176–12189.

90. Wang, M., Foster, D. J., and Pfeiffer, B. E. (2020). Alternating sequences of future and past behavior encoded within hippocampal theta oscillations. *Science*, 370(6513):247–250.

91. Blum, K. I. and Abbott, L. F. (1996). A model of spatial map formation in the hippocampus of the rat. *Neural computation*, 8(1):85–93.

92. Jensen, O. and Lisman, J. E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in neurosciences*, 28(2):67–72.

93. Mehta, M. R., Barnes, C. A., and McNaughton, B. L. (1997). Experience-dependent, asymmetric expansion of hippocampal place fields. *Proceedings of the National Academy of Sciences*, 94(16):8918–8921.

94. Redish, A. D. and Touretzky, D. S. (1998). The role of the hippocampus in solving the Morris water maze. *Neural computation*, 10(1):73–111.

95. Wikenheiser, A. M. and Redish, A. D. (2015). Hippocampal theta sequences reflect current goals. *Nature neuroscience*, 18(2):289–294.

96. Lisman, J. and Redish, A. D. (2009). Prediction, sequences and the hippocampus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521):1193–1201.

97. Euston, D. R., Tatsuno, M., and McNaughton, B. L. (2007). Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science*, 318(5853):1147–1150.

98. O'Neill, J., Boccara, C., Stella, F., Schoenenberger, P., and Csicsvari, J. (2017). Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. *Science*, 355(6321):184–188.

99. Vaz, A. P., Wittig Jr, J. H., Inati, S. K., and Zaghloul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. *Science*, 367(6482):1131–1134.

100. Wittkuhn, L. and Schuck, N. W. (2021). Dynamics of fmri patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nature communications*, 12(1):1795.

101. Denovellis, E. L., Gillespie, A. K., Coulter, M. E., Sosa, M., Chung, J. E., Eden, U. T., and Frank, L. M. (2021). Hippocampal replay of experience at real-world speeds. *Elife*, 10:e64505.

102. Yamamoto, J. and Tonegawa, S. (2017). Direct medial entorhinal cortex input to hippocampal CA1 is crucial for extended quiet awake replay. *Neuron*, 96(1):217–227.

103. Ólafsdóttir, H. F., Carpenter, F., and Barry, C. (2017). Task demands predict a dynamic switch in the content of awake hippocampal replay. *Neuron*, 96(4):925–935.

104. Ecker, A., Bagi, B., Vértes, E., Steinbach-Németh, O., Karlócai, M. R., Papp, O. I., Miklós, I., Norbert Hájos, T. F. F., Gulyás, A. I., and Káli, S. (2022). Hippocampal sharp wave-ripples and the associated sequence replay emerge from structured synaptic interactions in a network model of area ca3. *Elife*, 11:e71850.

105. Hasselmo, M. E. (2008). Temporally structured replay of neural activity in a model of entorhinal cortex, hippocampus and postsubiculum. *European Journal of Neuroscience*, 28(7):1301–1315.

106. Levenstein, D., Efremov, A., Eyono, R. H., Peyrache, A., and Richards, B. (2024). Sequential predictive learning is a unifying theory for hippocampal representation and replay. *bioRxiv*, 2024-04.

107. Shen, B. and McNaughton, B. L. (1996). Modeling the spontaneous reactivation of experience-specific hippocampal cell assembles during sleep. *Hippocampus*, 6(6):685–692.

108. Spens, E. and Burgess, N. (2023). A generative model of memory construction and consolidation. *bioRxiv*.

109. Barry, D. N. and Love, B. C. (2023). A neural network account of memory replay and knowledge consolidation. *Cerebral Cortex*, 33(1):83–95.

110. Káli, S. and Dayan, P. (2004). Off-line replay maintains declarative memories in a model of hippocampal-neocortical interactions. *Nature neuroscience*, 7(3):286–294.

111. Santoro, A., Frankland, P. W., and Richards, B. A. (2016). Memory transformation enhances reinforcement learning in dynamic environments. *Journal of Neuroscience*, 36(48):12228–12242.

112. González, O. C., Sokolov, Y., Krishnan, G. P., Delanois, J. E., and Bazhenov, M. (2020). Can sleep protect memories from catastrophic forgetting? *Elife*, 9, e51005.

113. Milstein, A. D., Tran, S., Ng, G., and Soltesz, I. (2023). Offline memory replay in recurrent neuronal networks emerges from constraints on online dynamics. *The Journal of Physiology*, 601(15):3241–3264.

114. Antony, J. W., Ferreira, C. S., Norman, K. A., and Wimber, M. (2017). Retrieval as a fast route to memory consolidation. *Trends in cognitive sciences*, 21(8):573–576.

115. Pu, Y., Kong, X.-Z., Ranganath, C., and Melloni, L. (2022). Event boundaries shape temporal organization of memory by resetting temporal context. *Nature communications*, 13(1):622.