REVIEW ARTICLE

# The evolving view of replay and its functions in wake and sleep

Graham Findlay[1,2,*], Giulio Tononi[1] and Chiara Cirelli[1,*]

[1]Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, USA and [2]Neuroscience Training Program, University of Wisconsin-Madison, Madison, WI, USA

*Corresponding authors. Graham Findlay, Department of Psychiatry, University of Wisconsin-Madison, 6001 Research Park Blvd, Madison, WI 53719, USA. Email: gfindlay@wisc.edu; Chiara Cirelli, Department of Psychiatry, University of Wisconsin-Madison, 6001 Research Park Blvd, Madison, WI 53719, USA. Email: ccirelli@wisc.edu

## Abstract

The term hippocampal replay originally referred to the temporally compressed reinstantiation, during rest, of sequential neural activity observed during prior active wake. Since its description in the 1990s, hippocampal replay has often been viewed as the key mechanism by which a memory trace is repeatedly rehearsed at high speeds during sleep and gradually transferred to neocortical circuits. However, the methods used to measure the occurrence of replay remain debated, and it is now clear that the underlying neural events are considerably more complicated than the traditional narratives had suggested. "Replay-like" activity happens during wake, can play out in reverse order, may represent trajectories never taken by the animal, and may have additional functions beyond memory consolidation, from learning values and solving the problem of credit assignment to decision-making and planning. Still, we know little about the role of replay in cognition, and to what extent it differs between wake and sleep. This may soon change, however, because decades-long efforts to explain replay in terms of reinforcement learning (RL) have started to yield testable predictions and possible explanations for a diverse set of observations. Here, we (1) survey the diverse features of replay, focusing especially on the latest findings; (2) discuss recent attempts at unifying disparate experimental results and putatively different cognitive functions under the banner of RL; (3) discuss methodological issues and theoretical biases that impede progress or may warrant a partial revaluation of the current literature, and finally; (4) highlight areas of considerable uncertainty and promising avenues of inquiry.

### Statement of Significance

Since its discovery, hippocampal replay during sleep has been assigned an important role in memory consolidation and integration. Here, we discuss how this role is evolving in light of the more recently discovered awake replay, which has several surprising features and many putative functions, from memory formation to retrieval and planning. We draw attention to the growing complexity of replay across the sleep/wake cycle and discuss recent attempts to provide a unifying explanation for this phenomenon.

## The Problem of Defining Replay

It is important to start by stressing the fact that there is no consistent technical definition of either reactivation or replay, and only recently the need to reach consensus in the use of these terms has been recognized [1]. In keeping with most of the literature, we use "reactivation" to refer to when a zero-lag measure of population activity computed over an epoch of 50–200 ms, for instance cell-pair spike correlations, is similar to a reference period. In some cases, the magnitude of similarity can provide a measure of "reactivation strength" that, when performed on consecutive epochs, will yield a timeseries. However, this measure does not take into account the temporal evolution of neural activity during each epoch. With the exception of the first seminal study [2], reactivation is always assessed in more than one cell, often between cell pairs, and provides a measure of cofiring within a variable time window of up to a few hundred milliseconds, but without any assumption regarding the order in which cells are reactivated. "Replay," in contrast, implies a measure that more directly assesses the temporal structure of activity in a population. In its weakest form, this might mean assessing whether population state transition probabilities are preserved, or checking that an asymmetry of spike cross-correlograms associated with an experience is reinstantiated. In its strictest form, replay implies that an extended sequence of neural activity—ideally firing of single units—matches a template. Often, this might mean correlating two spike trains or checking that the rank order of spiking across cells in a population is preserved. Also very popular is a Bayesian approach, in which a decoder is trained during active wake to estimate the animal's position from a brief (usually 10–50 ms) snapshot of population activity, and then applied later to candidate replay events to yield a posterior probability of position versus time that must match the reference period.

With the exception of some reactivation studies that have significant historical importance for the replay field, we have generally tried to focus here on sequence replay derived from single-unit data, but we do not always specify whether the template was a posterior probability matrix, a spike train, a rank-order sequence, etc. It would be impossible to review the literature this way, given the wide range of methods in use and the fact that no two studies—even from the same lead author—seem to share precisely the same methods. Combined with statistical and conceptual issues that make the study of replay particularly challenging (see Box 1), this lack of methodological standardization is one of the most problematic aspects of the replay literature, as it often makes comparing even congruent results very difficult. Unless explicitly stated otherwise (e.g. in the section on human replay), all studies mentioned herein were performed in rodents—mostly rats.

## The Classical Picture of Replay (1989–2005): The Link With Sleep, Sharp-Wave Ripples, and Memory Consolidation

The first report of a hippocampal cell being reactivated away from a stimulus to which it was tuned was provided in 1989 by Pavlides and Winson [2], who studied pairs of cells with nonoverlapping place fields recorded from freely moving rats. Their stated goal was to test the hypothesis that after exposure to a specific environment, the firing of place cells during sleep may represent "some form of memory processing." In line with this idea, the authors found that confining the animal to a given cell's place field elicited higher firing rates and increased bursting from that cell during subsequent nonrapid eye movement (NREM) and rapid eye movement (REM) sleep, but not during postexposure wake, compared to place cells that were not activated during prior wake. As mentioned, in this seminal study the firing pattern was studied separately in each cell, without any measure of cofiring. In 1994, after population recordings became feasible, Wilson and McNaughton found that pairs of cells whose place fields overlapped during behavior had an increased tendency to fire together during subsequent sleep [12], although this and other studies [13–15] actually failed to replicate the previous finding [2] regarding higher firing rates during postexposure sleep. The 1994 study suggested for the first time that entire assemblies were being coherently reactivated during postexposure periods of sleep, defined based on the presence of sharp-wave ripples (SPW-Rs) in the hippocampal EEG, and was followed a few years later, in 1996, by the first report that this reactivation preserved the order in which place cells fired during spatial navigation [15]. Over the next few years, increasingly compelling demonstrations of sequence replay over tens of cells were made during slow-wave sleep [14, 16]. This replay happens over timescales of up to 200 ms, much faster than the original behavior that spans several seconds.

SPW-Rs are one of the most distinctive patterns in the hippocampal local field potential (LFP) [17]. The sharp-wave component of SPW-Rs is a negative wave lasting about 40–100 ms that is prominent in the CA1 stratum radiatum. It is thought to reflect the synchronous depolarization of CA1 apical dendrites by synchronously bursting pyramidal cells in upstream CA3, where strong recurrent connectivity allows small numbers of excitatory cells to provoke rapidly spreading population bursts. Often, the sharp wave-induced excitation in CA1 also engages interneurons, and the rapid oscillation of interneuron-coordinated pyramidal cells generates a coincident "ripple" in the LFP at 120–200 Hz [18].

Although replay and reactivation events in rodents are closely associated with spiking during these ripples, there does not appear to be a 1-to-1 relationship between ripples and either replay or reactivation events. Most studies restrict their analyses to ripples or SPW-Rs, but when the detection of candidate events is not conditioned on ripple events, replay and reactivation are still found outside of ripples [13, 19, 20]. Whether or not every ripple has replay content is a much more difficult question to answer, since it is always possible that candidate replays that do not reach significance are, in fact, replays of contexts or episodes that were missed by the experimenter, or of events whose representations have transformed over time to the point of being unrecognizable.

There are many reasons why the close relationship between replay and SPW-Rs is significant, starting from the ability of SPW-Rs to effect wide-ranging patterns of activity in both neocortex [21] and subcortical structures [22]. Thus, the first observations of sequence replay immediately suggested a mechanism for memory consolidation (reviewed in Klinzing *et al.* [23])—in particular a "two-stage process" in which "information encoded in the hippocampus during awake, exploratory behavior is "replayed" fast during subsequent sleep [18, 24, 25]. Although there were early suggestions that reactivations were not tied to sleep per se [12, 13], the many connections between sleep and memory consolidation, in the absence of any reports of sequence replay during wake until much later [26], facilitated the view of replay

## Box 1. Methodological, Statistical, and Conceptual Challenges in the Study of Replay

The statistical issues with reactivation and replay analyses are exceptionally thorny, and with replay especially the choice of method can significantly affect which events are detected or reach significance (for excellent recent reviews, see van der Meer *et al*. [3] and Tingley and Peyrache [4]). In particular, for any given study, it is unclear how confident we should be in the chance distributions against which replay events reach significance. Replayed sequences are virtually never perfect matches to a template. Thus, rather than count the number of unambiguous matches over the course of a night and evaluate whether that number exceeds some threshold required for significance, a determination must be made that each individual event significantly matches the template. In order to do this, some kind of within-event shuffle—whether it be the identity of cells, the position of each cell's place field, the time of a decoded posterior probability over position, etc.—needs to underlie the construction of a chance distribution. The core problem is that there is no way to do this such that the factor of interest—the sequential order of spiking—is randomized while other parameters, such as firing rates, cell-pair correlations, place field shapes, etc., are preserved (see, for example, Fig. 4 of Foster's excellent review [5]). Thus, every approach increases the likelihood of spurious significance beyond what is necessary (i.e. what would be observed in random data), and our incomplete understanding of the structure of neural activity makes the estimation of false-positive rates difficult. Even if we had complete knowledge of this structure, virtually all methods have a false-positive rate that is heavily dependent on factors like the firing rates of the cells recorded (e.g. fast-firing neurons are more likely to fire first [6], and correlation between two spike trains is a function of spike rate, regardless of shared inputs [7]), the number of cells recorded, and the length of the sequence of interest. For example, one can check with a few lines of code that random rank-order sequences of length 10 will yield significant correlations with a random template up to 12% of the time, and this false-positive rate increases as the sequence length drops [4]. In the face of this, one always has to look at the proportion of all events that come out significant and compare that proportion across conditions (e.g. pre-task sleep and post-task sleep). When one wants to perform "second-order analyses" like comparing strength of replay across conditions, or essentially any analysis for which it is no longer sufficient to determine that activity is structured relative to some notion of chance, complexity explodes [3]. Of course, researchers in this field tend to be hyper-aware of these issues and use state-of-the-art, large-scale recording methods that offer high statistical power. Most analyses will impose many strict constraints on replay, and it is becoming increasingly common to see results reported for a wide range of parameter choices. Some studies are meticulous in their methodology and provide excellent examples of how to conduct and discuss second- and first-order analyses (for example, see Carey *et al*. [8] and Silva *et al*. [9], respectively). Nevertheless, the fact remains that we have limited understanding of how good our notion of chance is when it comes to spike sequences. It is not clear that replay is a binary phenomenon either, rather than a graded one. Replay scores often follow continuous rather than bimodal distributions, and casting a wider net to include more candidate replay events can substantially change the picture of replay [10, 11]. As mentioned, the literature covers a continuum of phenomena ranging from 1-step population state transitions that occur above chance to extremely strict single-unit spike sequences across tens of cells, to say nothing of the largely overlapping literature on reactivation and other measures of zero-lag population structure. What exactly the relationship is between reactivation and replay needs further attention. Perhaps reactivations encode a context or environment, and replays encode a specific experience or trajectory within that environment. Or, perhaps, reactivation and replay are just two faces of the same graded phenomenon.

as a sleep process. According to this view (simplified considerably), replay is a mechanism by which a hippocampally dependent memory trace is repeatedly rehearsed at high speeds during sleep, and in the process is gradually transferred to neocortical circuits. The idea that hippocampal place cells more-or-less faithfully recapitulate sped-up sequences of activity observed during prior awake behavior, viewed in the context of the two-stage model of memory consolidation, could be called the "classical" picture of replay [5]. Some variations of this model stress the fact that the transfer may never be complete, leaving the memory always, to some extent, dependent on the hippocampus, and/or emphasize that the transfer is associated with a cortical transformation of the original memory trace (e.g. gist extraction) [23, 27, 28]. In all cases, replay during sleep is seen as key for sleep-dependent memory consolidation.

## Replay-Like Activity During Wake: Immediate, Often in Reverse, and Occasionally Nonlocal

Some of the data in the original paper by Pavlides and Winson [2] suggested that the reactivations were strongest during sleep, but there was nothing in any of the early studies to suggest that reactivation or replay could not occur during wake. Indeed, Kudrimoti *et al*. [13] found that reactivations similar to those observed in 12 were more closely associated with the presence of SPW-Rs—the dominant pattern in the hippocampal EEG of rodents during slow-wave sleep—than with sleep per se. When the authors tested this by comparing post-task reactivations during sleep with quiet wake—when SPW-Rs are less frequent but still prevalent—they found that the two states were comparable. Although the focus on sleep persisted for years after this study, compelling sequence replay was eventually observed during wake [26].

The finding of sequence replay during wake immediately challenged the idea of sleep replay as a faithful recapitulation of prior well-rehearsed experience, because it found that these sequences could play out immediately after a single brief experience (the first lap in a novel track), and often in reverse [26]. In open-field environments, the tendency for place fields to be omnidirectional (i.e. true place fields, not sensitive to the heading of the animal) makes it difficult to say without doubt that a reverse sequence is not in fact a return trajectory played forwards. However, one can either exploit the fact that the animal is unlikely to run a return trajectory, or use

pre-experience sleep as a control to establish that reverse reply is not in fact "forward preplay" of a soon-to-be experienced trajectory [9, 29]. In 1D settings, the confound of running the return trajectory is greater, but the directional nature of place fields here can be exploited to disambiguate reverse replay from forward replay of the return [9, 26]. Even more compellingly, reverse replay can be seen in 1D environments when the animal only ever runs in one direction [30].

In many ways, the statistics of forward and reverse replay are largely similar. There do not appear to be any reports, for instance, suggesting that one is faster than the other, or covers longer spatial trajectories. During wake, both tend to begin at the animal's location [31–33]. However, there are clear examples (see e.g. Davidson *et al.* [34] and especially Gupta *et al.* [30]) of both forward and reverse replay of the immediate environment that start away from the subject. Despite these similarities, there are consistent and striking differences between forward and reverse replay in their sensitivity to reward and reward context. When animals are asked to repeatedly run linear tracks for water rewards, forward replay dominates before the run (~95% of replay), while reverse replay dominates after the run (~85%) [31]. The rate of reverse replay increases when an animal encounters an increased reward but decreases when the expected reward is reduced or removed altogether. In contrast, the change in reward does not appear to affect the rate of forward replay [35, 36]. When rewards are placed at both ends of a linear track, and the reward on only one side is increased, the rate of reverse replay increases on the side with increased reward, as expected, but also decreases on the other side of the track with unchanged reward, suggesting that reverse replay is sensitive to counterfactual scenarios or, at least, the amount of reward received relative to the total available [35]. Altogether, the strong modulation by reward has prompted the suggestion that reverse replay is a consolidation mechanism for learning values and solving the problem of credit assignment [5, 27] (see section on reinforcement learning).

Although the rate of forward replay does not appear sensitive to changes in reinforcing rewards, it is plausible that it is sensitive to changes in punishment. When rats are asked to run back-and-forth on a linear track and are given an electric shock at one end of the track, replay becomes biased toward the area where the shock was received, even though the agent avoids that area [37]. Unfortunately, the majority of place fields in this particular study were omnidirectional (~70%), so the authors were unable to discern forward from reverse replay. To date, there is very little literature examining the effect of punishment on replay.

For many years, it was assumed that replayed trajectories during wake at least reflected recent experience, even if they were not always highly faithful recapitulations of it. This fit with a theory that awake replays (especially reverse replay) might be related to residual or trace activity of recently active cells [29, 38]. It is now well established, however, that there is often replay of locations in the immediate environment that have not been recently visited (i.e. within the past 10 minutes) [8, 30]. The proportion of replays that are nonlocal in this way is as high as 40% in some studies, though such proportions are rarely reported. Proportions of forward and reverse events under these circumstances appear to be roughly similar to replay of recently experienced trajectories. Furthermore, "remote replay" during wake of entirely separate environments is possible. In the first study to report remote replays during wake [39], these events were just as frequent as local replay events, though the authors did not distinguish between forward and reverse replay. To our knowledge, no study has explicitly compared the proportion of forward and reverse events in remote replay as compared to local replay, and this proportion would likely be confounded by other factors present in the local environment, such as the availability of immediate reward.

## An Even Broader View: Replay as a Model of the World

### Model-based recombination of past trajectories

As its name implies, replay has been traditionally framed as the temporally compressed reexpression of some behavioral episode. However, replay can represent trajectories never taken by the animal. For example, in an open-field foraging experiment with systematically varying start and goal locations, replay was found to start at the subject and proceed toward the goal, even for entirely novel trajectories untraversed by the animal [20]. In another experiment where rats were placed into a T-maze with opposing arms blocked off so that they could see rewards in either arm but not reach them, formation of novel replay trajectories that led into rewarded arms was observed [33]. Novel forward trajectories have also been observed in several other studies [10, 30, 34]. These results are more compatible with replay as the spontaneous exploration of network attractor states, or the expression of structural knowledge about paths through the environment triggered by recent behavior.

Importantly, novel sequences cannot be constructed by so called "model-free" learning mechanisms, and heavily imply the existence of a cognitive map. In one of the more direct attempts to assess whether replay explores a model of the environment [40], rats were placed into an unusually large Y maze and run from the end of one arm to another, through the central choice point. It had been shown previously that in large environments multiple consecutive replay events (corresponding to separate ripple events, discussed later) could be "stitched together" to replay an extended trajectory [34]. In the forked maze, arm lengths were long enough that replay of a single run was likely to split across events. Interestingly, replay events tended to be segmented by the maze's bifurcation point, even when the replayed arms were of unequal length. Furthermore, changes in replay direction (e.g. from reverse to forward) were more likely to occur at this point. Thus, it seems that the maze's topology had some influence on replay structure. The authors also found that replay events of different trajectories covering overlapping parts of the maze recruited the same populations of cells (~94% overlap) to represent shared maze segments. Given that up to 50% of all place cells may be active in an environment, and a single place field is often covered by many different place cells, this degree of overlap was remarkable and suggested that the different trajectories might have been tied to a common underlying model, perhaps of the environment. An even more extreme suggestion that replay does not simply represent episodic content comes from a recent study by Stella *et al.* [10] that used an unorthodox pellet-chasing task to prevent the animals from following stereotyped paths, while also managing to keep their motion highly nonrandom. The authors found that in subsequent rest replay trajectories resembling random, Brownian motion tiled

the entire environment at several spatial and temporal scales, as if representing its overall statistics or traversable paths.

### Prospective replay and its role in planning

An open question is whether "prospective replay"—that is, replay which is predictive of the animal's future trajectory—is related to planning. It has often been suggested that forward replay and reverse replay might serve prospective and retrospective roles, respectively [31, 41]. One of the strongest pieces of evidence in favor of this interpretation comes from a study by Pfeiffer and Foster [20] that found that replays just before navigation toward known, hidden goals were biased toward these goals, and could even predict the animal's movements after finding the goal, but did not predict the animal's immediate movements during random foraging. Another recent study found that forward replay of a T-maze arm—experienced visually yet not explored—was a prerequisite for the correct (i.e. above-chance) choice of that arm. More generally, it was found that replay prior to navigation was biased toward the correct arm, although it remains unclear whether the effect was driven by planning rather than prior sighting of the reward [33]. A recent study that observed replays in a radial maze [42] found that during reference memory tasks, replays at the choice point were more likely to be forward replays that predicted the animal's path, while during a working memory task, choice-point replays were more likely to be reversed and of arms that the animal had already visited. Another recent study [41] reported a strong link between forward replay and upcoming trajectories, and between reverse replay and completed paths—but not at choice points.

Meanwhile, other studies have reported cases of replay being biased *away* from the correct, preferred, or upcoming trajectory [8, 30, 40]. When rats were rewarded on only one arm of a continuous T-maze and ran continuous laps on that side of the track, they replayed the opposite arm even more than when they alternated between arms [30]. In an impressively meticulous recent study, rats were water restricted or food restricted on alternating days, and placed each day into a T-maze offering free choice between food and water in opposite arms [8]. The replay was heavily focused on the nonpreferred, nonexplored arm, and multiple analyses suggested that it was actually the animal's motivational state, rather than its immediate past or future experience, that drove this paradoxical effect. The authors speculated that this sort of replay might serve to preserve or protect in some way the options not taken, or might simply reflect the fact that in overtrained animals making easy decisions, replay circuitry might be freed up for other purposes not related to the task at hand. Consistent with this latter possibility, it was previously found [43] that reactivations ceased to predict upcoming paths after a certain point during learning.

### Preplay: a highly debated phenomenon

All the reports of novel trajectories described above find that they are composed of subpaths previously traversed by the animal or, at least, visually experienced and inferred. This makes sense if replays are constructed from learned information: how could a rat replay an environment of which it had no knowledge? This sort of novel sequential activity, though it may predict novel future experience, needs to be distinguished from a controversial phenomenon dubbed "preplay." Preplay refers to the sequential firing of place cells in a way that presages their relative ordering in a not-yet experienced environment. If putative replays might in fact be preplays which the experimenter cannot be aware of and control for, the entire premise and study of replay—even as a prospective computation performed on learned information—is threatened. The first study that reported preplay [44] recorded hippocampal firing sequences during periods of awake rest, alternating with periods of running, on a familiar track. The twist was that these recordings were followed by the exploration of a visually isolated novel arm contiguous with this familiar track. Interleaved with replays of the familiar arm, the authors claimed to see firing sequences that were later matched by exploration of the novel arm. Several additional reports of preplay, mostly from the same authors, have followed since [44–48]. Other authors [33] have used the term "preplay" to describe their findings, but the fact that animals had visual access to the small, preplayed area complicates this interpretation.

Taken at face value, these results are consistent with several recent observations about the remarkable low dimensionality of certain brain activity [11, 49], the brain's tendency to recycle existing neural patterns for new purposes [50], and more specifically with an attractive theory that the hippocampus maintains a reservoir of preconfigured spike sequences that are dynamically assigned to novel experiences [48, 51]. In general, it is not clear where to draw the line between replay/preplay and spontaneous exploration of a network's attractor states. In any case, the hypothesis offered to explain preplay is that internally generated, preconfigured spike sequences are formed around a backbone of rigid synapses with slight, experience-dependent adjustments provided by a second set of plastic synapses at runtime [47, 51]. On the other hand, other authors [10] have concluded that no such reservoir could realistically be large enough to support the Brownian trajectories that they observed, and well-powered attempts to directly replicate preplay findings have failed [9]. Furthermore, there are several reasons to suspect that preplay is the artifact of flawed analyses [5]. Put conservatively, "Virtually all investigators agree that replay in post-task sleep is stronger than preplay in sleep before the task, although exactly how much is still under investigation and debate" [3].

## Is Replay Fundamentally Different Between Wake and Sleep?

As summarized above, most recent experiments have focused on awake replay and unfortunately, very few studies have considered both wake and sleep. Thus, data on the differences in replay between the two behavioral states are limited. Of note, most studies that focused on sleep used scoring criteria that all but guarantee their data were contaminated by quiet wake. Some studies treated the entire post-task period as "sleep" or "rest," lumping immobility and sleep together [10, 52], or only performed scoring by visual assessment of posture [16]. Even when studies scored using stricter criteria—often using a theta:delta band power ratio applied to a single hippocampal field potential—much analysis was then performed after lumping sleep and quiet wake together [47].

There is some evidence that reverse replay is less frequent in slow-wave sleep than wake—both implicitly from a lack of reports of reverse replay during sleep, and explicitly using spike

cross-correlograms in Wikenheiser and Redish [53]. On the other hand, it is unclear that this is a feature of sleep per se, given that the rate of reverse replay declines both with experience in a given environment [26, 54] and with time [39]. It is also possible that sleep may contain the reactivation or replay of contexts other than the one of interest to the experimenter, which would be undetectable by traditional methods.

Another question is the extent to which replay during sleep is always a form of remote replay. Ironically, the seminal 1989 paper from Pavlides and Winson [2] is still one of the only examples of recording both wake and sleep in the same enclosure, and the authors were also careful to ensure that the animal did not sleep inside a recorded place field. In the great majority of studies, however, animals almost never sleep in the open fields or mazes where they undergo the behavioral experience and place-cell mapping of interest to the experimenters. Instead, their postrun sleep takes place in separate, very small enclosures that do not share visual cues with the environment of interest. Thus, as far as we are aware, all sleep replay in the literature is considered remote. Because it is not clear that local and remote replay have inherently different statistics during wake, it is difficult to make any predictions about what replay of a local environment ought to look like during sleep. Importantly, a population of cells in CA1 and CA2 do reflect a rat's immediate location during periods of hippocampal desynchronization without ripples that occupy roughly one fifth of total sleep [55]. CA2 cells may be capable of triggering SPW-Rs, and more so during NREM than wake [56], which raises the possibility that nesting location could influence these processes [57].

There is some direct evidence that reactivation and replay events during sleep are less frequent and/or weaker and less structured than during wake. In a recent study from Tang *et al.* using rats trained in a spatial alternation task, measures of cofiring revealed that the reactivation between pairs of CA1 excitatory neurons, and between CA1 and prefrontal cortical neurons, was stronger during awake SPW-Rs than during sleep SPW-Rs [57]. The replay of CA1 sequences also occurred more frequently during awake SPW-Rs than during sleep SPW-Rs, and measures of ensemble reactivation of the CA1–prefrontal network suggested stronger reactivation strength during wake than during sleep. Of note, the stronger and more structured reactivation during wake occurred despite the fact that awake SPW-Rs were shorter and of smaller amplitude than sleep SPW-Rs, and unlike during sleep they were not coordinated by the occurrence of cortical slow waves and spindles [57]. It is worth mentioning that this study followed the standard practice of recording sleep in a dedicated nest box, so the stronger awake reactivations may be local to the maze environment, whereas the weaker sleep replay may be remote. This local/remote distinction could plausibly contribute to some of the differences reported between wake and sleep, and future studies might explore this possibility. Another study from Grosmark and Buzsaki [47] showed that the mean "sequenceness" score of candidate events was lower in sleep than wake, which may be in line with the hypothesis that replay during sleep is more "disjoint" [58], or has lower fidelity to prior wake experience [39, 57]. It has been speculated that this comparatively noisy reactivation may reflect the active transformation of memory during sleep [59], or the simultaneous activation of multiple memories undergoing integration [60].

Putative demonstrations of hippocampal replay during REM sleep remain extremely rare and exhibit several odd features.

These replays do not appear to depend on SPW-Rs—which are very rare during REM sleep—and may not exhibit the "fast-forwarded" or temporally compressed quality seen in other states [61]. Moreover, most significant template matches were actually identified during the REM sleep episodes that occurred before the task, not afterwards, and therefore the term "replay" may need to be qualified [61]. One of the few other studies to examine REM sleep [13] failed to find significant place-cell reactivation during this phase.

In short, based on the limited data currently available, perhaps the most obvious difference between wake replay and sleep replay is the lack of modulation by slow waves and spindles during awake replay [57]. By design, sleep replay is almost always remote and wake replay is more often immediate, but it can be remote as well. Moreover, replay can be forward and reverse in both wake and sleep, but whether their relative frequency differs across behavioral states remains unclear. As we have seen, replay-like activity has been attributed several functions in addition to memory consolidation, from learning values to planning, but the specific contribution of sleep replay in these processes is unclear. For memory consolidation, one influential model, mainly based on in vitro and computational experiments, predicts that the high cholinergic activity of wake promotes encoding by enhancing afferent inputs to the hippocampus, while the low cholinergic tone of slow-wave sleep favors consolidation by exciting the CA3 association system, thus triggering SPW-Rs and their propagation outside the hippocampus [62]. Consistent with this hypothesis, medial entorhinal cortex (MEC) may exert greater influence over replays during wake than NREM, whereas CA3 input to CA1 is essential in either state [63]. However, because both wake replay and sleep replay promote memory consolidation, it is unclear whether replay during sleep is special [27]. If so, it may be because the levels of acetylcholine are expected to be lower in NREM sleep than in quiet wake [23]; this assumption, however, should be confirmed using methods that can detect extracellular levels of acetylcholine with high (1-s) temporal resolution [64]. Another suggestion is that despite being less frequent and weaker than in wake, replay during sleep is uniquely coordinated by slow waves and spindles, and thus is in a strong position to integrate multiple experiences across large thalamocortical and hippocampal networks [57].

## The Time-Course of Replay and Its Implications for the Memory Process

There is evidence that both the strength of replay and its speed change with time, with obvious implications for the role of replay in the memory process. However, the time-course of replay is surprisingly difficult to assess. Early studies found that the strength of reactivations declined rapidly during the first post-task hour [12, 13], and the same trend has been reported for post-task sequence replay [14, 19, 26]. Consequently, many experiments are now designed to focus on this brief post-task window, and largely ignore later periods. If replay decays this rapidly, however, it imposes a serious constraint on any hypothesis about its role in memory consolidation, as it ties associated processes to this critical window [23, 65]. It is possible but unproven that replay might resume at unobserved time points, continue as transformed sequences undetectable to the

experimenter, or persist at very low levels that make it difficult to observe.

A recent study from Giri *et al.* [66] sought to resolve this conundrum, observing that previous experiments examining the time-course of reactivation or replay had all taken place in relatively familiar environments, and proposing that novel environments might trigger more sustained effects. In fact, previous studies had not overlooked the possibility that novelty might significantly modulate reactivation time-course. Tatsuno *et al.* [67] had examined the effects of novel objects in a familiar environment and found no evidence of sustained reactivation or replay. In another experiment, Kudrimoti *et al.* [13] had trained rats to run one half of a figure eight maze (with the central arm shared, i.e. a "digital" eight), never allowing them to enter or see the second half. In a "novel environment" condition the animal first ran in the familiar half, then in the unfamiliar half, and then finally again in the familiar half, with the barriers being repositioned each time to occlude either the familiar or novel loop of track. Thus, this "novel" condition was arguably two parts familiar, and only one part novel. Part of the reason that this was done was to establish that there was no significant remapping between the beginning and end of the task, in order to argue that the same set of cells were being recorded in the novel and familiar conditions. Although reactivation in the novel condition appeared to decay more slowly than in the familiar-only condition and had a time constant outside the recording duration of the study, Kudrimoti and colleagues [13] saw significantly *less* reactivation of the novel condition in their data, never reaching even the lowest levels observed in the familiar-only condition.

Novelty, however, comes in many forms, and the early studies discussed above predated the finding that some forms of novelty induce "rate remapping" of existing place fields, while other forms of novelty induce "global remapping" that elicits a new hippocampal map [68]. It is this global remapping that is believed to cause sustained reactivation [66]. To test this hypothesis, the authors analyzed data from four different experiments with extended post-task recordings, in which both the task and environment were designed to be novel. Using the same methods as Kudrimoti [13], as well as a variant approach to better account for changes in reactivation due to changes in firing rate between pre- and post-task conditions, they observed reactivations extending several hours longer than previously reported. Moreover, their data suggest that reactivations may be present during REM sleep, peak in strength several hours after the novel experience, and be stronger in slow-wave sleep than REM sleep or wake, superficially at odds with other data [13, 57]. Meanwhile, their control analysis of a familiar task in a familiar environment shows an extremely rapid decay in reactivation strength, as expected.

While these are intriguing results [66], their implications for replay proper are unclear. Although the authors used temporal bias in spike cross-correlograms as in a previous study [15] to demonstrate that reactivations preserved some of the order in which cells fire during the task, this constitutes replay only in a weak sense. Given the key role assigned to replay and/or reactivation as mechanisms for active systems consolidation [23] and/or contextual binding [69], it is critical that future work focuses on replicating and refining these these very interesting findings.

The dynamics of replay formation are easier to summarize. The older, aforementioned studies that failed to find significant reactivation after certain kinds of novel experience suggested that many repetitions of an experience were required to elicit reactivation, supporting the idea that reactivation reflected well-learned content. However, this conflicts with newer, reliable reports that even single experiences can spawn multiple replay events (see e.g. Foster and Wilson [26]). Indeed, the same study which demonstrated that sufficiently novel conditions might result in extended reactivation also suggests that even the first minute of a novel task session can be detected as repeated reactivations after the task [66]. Thus, it seems that replay and reactivation can be induced very rapidly under appropriate circumstances. The speed of replay also appears to change with time. The first report of awake replay found that the strength of replay increased across consecutive replay events, each separated by a new experience (one lap), while the speed of replay declined across laps [26]. Why exactly this decrease in replay speed occurs is an open question [5].

## Existing Tools for Manipulating Replay

Even if replay can predict an animal's upcoming trajectory under certain conditions, it does not follow that this activity is necessary for decision-making. Such a claim would have to be supported by intervention that disrupts the replays in question. Unfortunately, nearly every effort to disrupt replay is based on detecting and disrupting SPW-Rs rather than replay sequences per se, confounding interpretation in these terms. For example, Jadhav *et al.* [70] used electrical stimulation to disrupt SPW-Rs during an E-maze task that required both reference memory and working memory, and found that the animal's choice performance on the working memory component was impaired. In a rare gain-of-function study, Fernández-Ruiz *et al.* [71] used optogenetic stimulation to prolong ripples, and saw a boost to spatial working memory performance. Similar manipulations during sleep have also pointed to a link between SPW-Rs and task performance [72–75]. While important, these studies also illustrate one of the largest issues facing replay researchers: there are no good tools for precisely manipulating specific sequences of activity. Recently, Joo and Frank [27] reviewed many studies that used SPW-R manipulations to evaluate the evidence for functional subtypes of SPW-Rs (e.g. SPW-Rs for memory consolidation versus SPW-Rs for planning, similar to the proposal that forward and reverse replay form functional types). They concluded that there is little evidence for such a distinction. Rather, they propose that each SPW-R serves multiple cognitive functions, and that many of these seemingly disparate functions may in fact reflect a single, underlying process.

The most sophisticated attempt at replay-specific manipulation comes from a study by Gridchyn *et al.* [52] that used online decoding of ensemble spike patterns [76] to detect reactivation of two different open-field cheeseboard environments during "high synchrony events" (defined at the spike-level, rather than in the LFP, and usually preceding SPW-Rs by ~50 ms). Rats were trained to locate hidden food rewards in each environment, and any high synchrony event not confidently related to the control environment was selectively disrupted during sleep using optogenetic inhibition of pyramidal neurons in CA1 using the neural silencer ArchT. The primary finding, as expected, was that spatial memory recall in the target environment relative to the control environment was impaired. This effect could have been due to the disruption of place field map formation, but because their intervention only transiently affected place field maps and

continued to affect performance after maps had restabilized, the authors argue that the main consequence of their manipulation was to interfere with the proper association of the cognitive map and its context. Perturbations in this experiment were targeted at periods of sleep, and it is unclear whether similar perturbations could currently be justified in wake, where it seems more plausible that basic processes of cognitive map formation would be disrupted and confound the interpretation of results [77–79]. Of course, despite being selectively triggered, even this state-of-the-art manipulation was much less selective in its effects: the entire hippocampus was perturbed every time the target context reactivated, more excitatory cells were disinhibited than inhibited, inhibited cells exhibited significant rebound firing, and overall ripple power was affected. Now that we are entering an age where optical detection and perturbation of replayed sequences at the unit level may soon be possible [80], we will likely see these methods combined with spatial navigation virtual reality tasks [81] to finally yield targeted, precise perturbations of replay.

## Human Replay of Nonspatial Task Elements

Although invasive studies only possible in rodents offer one of the most promising paths forward for the field, another frontier is opening up in the form of human replay. Far from merely suggesting that the vast rodent replay literature has cross-species relevance, these studies are allowing experimenters to divorce the study of replay from spatial navigation, and suggest ways in which replay might serve more general cognitive functions.

Although reinstated patterns of neural activity from wake experience have been seen many times over in humans, especially in the memory retrieval literature, very few of these bear close similarity to rodent replay. They may not, for example, occur predominantly during ripples, have compressed timescales, show context-dependence on reward, have strongly sequential structure, occur spontaneously during post-task rest periods, and so on [82].

The most compelling demonstration of human replay comes from a recent magnetoencephalography (MEG) study by Liu *et al.* that trained classifiers to detect item-specific patterns of activation associated with visual images, and then decoded the sequential reactivation of these patterns during pre- and post-task periods of rest [83]. These periods lasted 5 min and occurred between two learning phases of the experiment. In the first of two experiments, participants were shown a sequence of images and told that this actually represented two separate sequences interleaved and out-of-order. Thus, two true sequences [WXYZ] and [W′X′Y′Z′] might be presented as [YZY′Z′][WXW′X′]. Participants were explained the rule mapping the observed sequences onto the true sequences, and then returned the next day to apply it to a new set of images. In post-task rest, the "true" sequences were replayed, mostly forward over ~200 ms, coincident with ripple-band power increases source localized to the medial temporal lobe. Interestingly, the visually experienced sequence was not replayed. At the end of the second day, participants were told that one true sequence was associated with a reward, then told that they would shortly be asked if randomly presented images belonged to the rewarding sequence. During rest immediately after being given the rewarding sequence but before testing, replay of true sequences was again observed, but this time in reverse.

In a second experiment, designed to reduce the likelihood that replay of visually experienced sequences could masquerade as replay of true sequences, all 1-step structure that existed in true sequences was removed from visual sequences. For example, if [WXYZ] and [W′X′Y′Z′] were true sequences, the order of visual presentation might be [Z′XY′Y...], such that W would never transition to X, X to Y, Y to Z, and so on. Replay of true sequences was observed as before, but this time the experimenters also asked whether the decoded activity included information about the true sequence that each item belonged to, as well as its position in that sequence, above and beyond any information about stimulus identity. For example, did items belonging to a particular sequence (e.g. sequence 1) all share representations absent in the other sequence (e.g. sequence 2), and did items in both sequences belonging to the same particular position (e.g. position 4) share representations absent in the other positions (e.g. positions 1–3)? Indeed, they found detectable neural representations of sequence membership and position that activated ~50 ms prior to the representation of each stimulus' identity, suggesting that these abstract, "factorized" neural codes facilitate retrieval of the correct image. Finally, the authors looked at pre-task rest, before images were ever shown. They found that they could detect reverse replay of position codes, but not stimulus identity. This would have constituted true preplay, as the participants were familiar with the task (and therefore its sequence and position aspects) from training the previous day, but the precise images in question had not been experienced yet. Consequently, they called this phenomenon "transfer replay," suspecting that it might be involved in generalizing previously learned rules to upcoming experience.

These results, speculatively applied back to the rodent literature, lend strong support to the idea that replay is not a recapitulation of experience, but rather represents a model of the world. They also fit well with prior findings that the hippocampal system in humans serves an abstract or conceptual—rather than a necessarily spatial—cognitive map [84–86]. It is consistent with prior MEG studies showing fast reverse reactivation of visual sequences [87], and with a recent fMRI study that reported "replay" (actually a very different but intriguing sort of structured, post-task reactivation) of abstract, nonspatial task states [88]. Unfortunately, there is still no evidence of human or nonhuman primate replay using single-unit activity from intracranial electrodes, despite the fact that place-like cells [89], concept cells [85], ripples [90], and reinstantiation of sequential spiking activity [91] have all been observed in intracranial data.

## Replay Beyond the Hippocampus

One of the more intriguing aspects of replay observed using MEG is the fact that most of the decoded activity is probably cortical. There is not enough space here to review the cortical–hippocampal dialog with respect to memory reactivation, but we can ask two questions of more limited scope: First, does replay as it is defined in CA1 exist outside of the hippocampus? Second, what is the rest of the brain doing just before, during, and after hippocampal replay?

### Cortical replay

In order to answer the first question, Ji and Wilson [19] recorded spiking activity in the primary visual cortex and CA1 as rats ran alternating trajectories around a figure-8 maze. In addition to classical place cells in CA1, the authors observed

that certain cells in V1 also exhibited spatially localized firing. It was assumed, especially given the lack of direct input to V1 from the hippocampus, that these cells were being driven by visual stimuli that were specific to particular locations on the track. When rank ordered by time of peak firing on a lap, these V1 cells offered repeatable sequences that could be template matched to later activity, akin to regular place cells. In post-task slow-wave sleep, these sequences were found to reoccur during "frames"—essentially UP states, but defined at the level of multiunit activity. Cortical replays and hippocampal replays tended to be temporally coincident—perhaps unsurprising given the tendency of SPW-Rs to follow cortical DOWN–UP transitions [92]. Although multiple analyses suggested that cortical and hippocampal replays of the same trajectory were coincident above chance, the small number of such events detected (only 9 in post-task sleep) made it impossible to determine whether cortical or hippocampal replay reliably preceded the other. Since this report [19], something akin to replay has been reported during slow-wave sleep in cingulate and prelimbic cortices [41, 93], parietal cortex [94], and possibly even ventral striatum [95]. The relationship of these events to each other, to replay in CA1, and even to SPW-Rs is very much unclear.

An open question is what kinds of analyses are appropriate when assessing replay in cells whose tuning functions differ significantly from unitary place fields. For example, one study [96] reported that grid cells in the deep layers of MEC (which receive hippocampal output) exhibited replay coordinated with place cells in CA1, while another [97] reported that grid cells in the superficial layers of MEC (which provide hippocampal input) also exhibit replay, but independently of CA1. While it is entirely conceivable that these disparate findings reflect a real difference between superficial and deep entorhinal cortex, it is also likely that the first study did not adequately control for the multiple receptive fields of grid cells, increasing the odds both of spurious replay detection and spurious coordination [98]. As long as we are only interested in replay of spike sequences, this is not much of an issue: If a cell responds to stimuli (e.g. locations) A and B, and fires during a replay, what does it matter if the cause was upstream activity related to A vs B? But if we care that specific mnemonic content is being replayed (as is implicitly assumed by many of the Bayesian decoding methods used for detecting replay in CA1), then care needs to be taken when interpreting putative replay in cortical cells whose tuning functions may be entirely unknown.

### Coordination of hippocampal replay with other structures

Perhaps the best attempt to situate hippocampal replay within the context of cortical activity comes from Rothschild *et al.* [99], who recorded spiking activity in the auditory cortex (AC) and CA1 of rats as they learned a sound-guided navigation task. During short (20 min) training-interleaved sleep episodes, the authors found transient strong reactivations of AC, and that population activity in AC during and up to 400 ms before SPW-Rs predicted ripple content, whereas ripple content predicted AC activity during and up to 400 ms after the start of a ripple. This result held even when activity in AC was biased by presenting several auditory cues during the sleep period. Although the task-relevant auditory cue appeared to elicit more SPW-Rs relative to other cues once the task was solidly learned, there was no additional evidence that the task-relevant cue or its associated AC activity patterns had any special relationship to hippocampal reactivation.

The suggestion of a cortical–hippocampal–cortical loop in primary sensory cortex has many intriguing implications, but it remains to be seen if this will generalize to other cortices. There are several suggestions that frontal and prefrontal regions—in particular primate ventromedial prefrontal cortex and rodent prelimbic cortex—may activate especially early with respect to SPW-Rs, hinting that they may play a role in initiating replay [27, 83, 100]. In general, a better understanding of how hippocampal replay relates to extra-hippocampal activity is missing from the literature. The impact of SPW-R-associated activity on putatively downstream regions is poorly understood, and literature linking replay to subcortical and thalamic activity is scarce [27]. Given the considerable complexity of the hippocampal–neocortical dialog during and around SPW-Rs [11, 21, 101, 102], the simple scenario of the hippocampus broadcasting uniform mnemonic content to several downstream regions may need revision.

## Attempts at Unification Under the Banner of Reinforcement Learning

Reinforcement learning (RL) is a mathematical framework for describing how internal representations of states and actions can help an agent to maximize cumulative long-term reward when learning solely from rewards and punishments by trial-and-error [103]. Based on the observation that biological agents are most likely to learn through something like RL [104], RL methods have been applied with great success in recent years to problems of artificial general intelligence that require flexible behavior in changing environments and, especially, the association of rewards with states or actions that may be removed in time and space from the reward itself [105]. In brief, RL amounts to propagating value between temporally or spatially adjacent states for the purposes of this "credit assignment."

When framed as a sequential decision problem, spatial navigation often requires an agent to perform credit assignment. For example, after finding a rewarding Cheerio in a maze, a rat who wants to refind that Cheerio on subsequent laps will need to connect it with a series of actions taken at maze choice points well before the reward was even anticipated—"How did I get here, and how can I get here again?" RL methods generally solve this problem by "planning backwards," propagating the reward received at a goal state (e.g. maze location) along a reverse sequence of intermediate states and actions, updating each with the knowledge that it might yield the reward if visited in the future. This kind of reward propagation, in which past states and actions are visited in reverse, hints at a similar role for reverse replay. However, if RL is going to explain replay, then it must account for forward replay too.

For many years, it has been suggested that forward and reverse replays might serve different purposes. For example, reverse replay might serve RL-based learning, and forward replay or theta sequences might serve model-based planning, presumably using the action policies formed by reverse replay [70, 106, 107]. Unfortunately, even back-of-the-envelope calculations argue that forward replay is unlikely to be a computation capable of considering competing action policies for immediate

action without leaving the animal mired in thought [5, 27]. For every action immediately available to an animal (e.g. "turn left" or "turn right"), there is a constantly branching tree of possible future actions that is infeasible to search through on the fly. In an excellent review, Joo and Frank [27] argue for an absence of SPW-R functional types and, implicitly, an absence of replay functional types. Their proposition, instead, is that the separate processes often attributed to different forms of replay are, in fact, different faces of one process, as yet poorly understood but likely encompassing decision-making, planning, recollection, imagination, and memory consolidation.

## A modern account of replay and RL

Despite early suggestions that place fields and replay might be involved in an RL process [26, 108], essentially no promising formalization of replay within the RL framework took place until very recently [58]. Specifically, Mattar and Daw propose a theory that conceives of a place cell's reactivation as a "Bellman backup"—the individual step of computation that uses a single past action (e.g. "turned left at choice point #1") and its outcome (e.g. "eventually found Cheerio") to improve future action choice. Naturally, learning from certain action–outcome pairs will improve future action choice more than others, and because each backup operation takes precious time, it makes sense to prioritize the order in which they should occur according to their expected value. Moreover, the expected utility of a backup might change depending on the current state and goal of an animal. By recognizing this and carefully formalizing the expected value of a given backup, Mattar and Daw propose a theory of which experiences should be considered, and when, during deliberation. Furthermore, by recognizing that association of a reward with a goal state can happen upon reward consumption, and subsequent backups can happen any time between that moment and future choice, they outline an effective way for learning to occur at points temporally far removed from the actual experience.

In brief, their theory revolves around prioritizing backups (i.e. which place field should be reactivated next) according to a product of "need" and "gain." Gain is essentially the answer to "how much more reward can I expect to earn by acting optimally from this location after performing this backup," while the need term is essentially "how likely am I to visit this location soon?" Because gain is closely linked to whether or not the animal will actually change its behavioral policy if a backup is performed (e.g. reducing the reward associated with the best possible outcome is not going to change anyone's behavioral policy if that outcome remains the best, and therefore generates zero gain), asymmetric effects of increasing versus decreasing reward values lead to very different effects on behavior than simple surprise-driven reward prediction error frameworks. Furthermore, because there is nothing in the model that limits replay trajectories to those continuously traversed by the animal (i.e. it can compose previously experienced trajectories to form novel ones), it can learn policies that would be missed by traditional temporal difference learning, which is essentially equivalent to performing backups in the order that faithfully recapitulates experience [109].

With this framework in hand, the authors are able to capture and explain a stunning array of empirical findings including: replay accelerates learning on simulated navigation tasks that can still be learned without replay [70]; forward replay dominates before a run, and reverse replay dominates after a run when reward is present at the end, but both can occur in either context [31]; most, but not all, replays begin at the agent's location [26, 30, 31]; replays are biased toward the agent's location, toward choice points [36], toward reward [31, 32], and the bias toward reward persists during remote replay [33]; replay can construct trajectories not previously traversed by the agent [20, 30]; the number of significant replay events in both directions decays with experience [26], as does individual place field reactivation [54]; and when events do occur, the probability that they include specific states increases with the number of visits to those states [38].

Perhaps most interesting and compelling is their ability to reproduce in simulations several empirical findings that relate to the asymmetric effects of prediction errors. For example, it is known that forward and reverse replay have different sensitivities to reward context: the frequency of reverse replay increases when rats counters a greater reward than expected, but actually decreases when the animal encounters a poorer reward than expected, and is sensitive to the amount of reward received relative to the total reward available (i.e. to alternative polices). On the other hand, the rate of forward replay does not change much, if at all, in either context [35, 36]. Finally, they are able to reproduce the finding that punishment (e.g. electric shock, framed here as negative reward) biases replay toward where the punishment was received, even though the agent avoids that area [37]. It should be noted that the authors of this study [37] were unable to determine replay direction due to a high proportion (~70%) of bidirectional place fields. Mattar and Daw [58] do not report if forward or reverse replay drove their simulated effect, but this could constitute an interesting prediction of their model.

This new framework is reminiscent of the proposition by Joo and Frank [27]—in their case made more on the basis of empirical findings than theoretical considerations, but entirely complementary—that the separate processes often attributed to different forms of replay are, in fact, different faces of one process. This new framework also paints a picture of replay as an active process, driving computation and change throughout systems, rather than simply "strengthening" memory traces. That is, memories are not things encoded independently of the need for action in an unbiased, dispassionate way. Forming a memory is, on this view, a sort of action. More speculatively, it is interesting to note that there is nothing in their model [58] that limits states to being spatial, and this could be taken as a prediction—now borne out—that we should see replay of nonspatial task states where we are sensitive to them [88].

On the other hand, there is just as much new literature that does not fit neatly into the Mattar and Daw framework [8, 10, 42]. Three important limitations of this model deserve mentioning, some of which are relevant for the sleep field. First, it does nothing to explain when replay should take place. During their simulations, in which agents traversed various 1- or 2D environments, agents were simply forced to perform backups between trials ("quiet wake"), and then the place field reactivations that occurred were subjected to the traditional methods used to detect statistically significant replay events. Thus, the form that replays took was guided by the model, but the simple fact of replay was not. This paradigm also assumes that events not detected as part of statistically significant replay were part of the same fundamental process (prioritized Bellman backups), which may not be the case.

Second, their gain term, roughly "how much more reward can I expect to earn in a given state after performing this backup," is something that the brain needs to compute, and there are many ways that this could be done. For the purposes of their simulations, the authors simply used the true gain, but this is only possible from the omniscient perspective of the experimenter. Although this tells us how theoretically optimal replay ought to behave according to their model, the brain would have to e.g. substitute past gain as a proxy for estimated future gain, or have some other heuristic. Future experiments could and should attempt to distinguish between candidate heuristics.

Finally, their model assumes that spatial navigation can be expressed as a Markovian decision problem, in which the optimal action in a given state is independent of state history. This is obviously not realistic in, for example, spatial working memory tasks. There are several ways that non-Markovian problems can be reframed as Markovian, for example by defining states such that they take into account trial memory. It may not be unreasonable to expect that place cells behave this way [110], but work still needs to be done to apply the Mattar and Daw framework to these kinds of tasks.

## Conclusions and Future Directions

The picture of replay that has emerged in recent years is one of a highly variable, model-based process whose most notable features include modulation by reward, engagement by nonrecent and nonlocal contexts, generativity, and a possible relevance for abstract (i.e. nonspatial, possibly nonmnemonic) computations. It is not entirely clear whether replay is better conceived of as a random process biased by experience, or something more algorithmic. The diversity of replay, while potentially reflecting wide-ranging and profound effects on several systems, also needs a unifying explanation. Alternatively, we may have to conclude that replay is not a unitary phenomenon.

Conceptually, one must ask whether there is anything special about the reactivation of readily interpretable place-cell sequences as compared to "nonsense" replay of place cells, or even of nonplace cells. Of course, studying replay of readily interpretable place-cell sequences is massively convenient. It is easy to design experiments that will elicit sequential activity in predictable ways; such sequences are easily tied to overt behaviors and rewards; spatial structure is practically guaranteed to be ethologically relevant, especially in rodents. Certainly, if one were to open the space of possible template sequences to include all spiking activity observed during active wake, analysis would become computationally and conceptually daunting. And yet it does seem that there is an opportunity, especially considering recently revived skepticism of representationalism [51], to consider the replay of neural patterns that might not be easily decoded or otherwise interpreted by the experimenter. We do not, after all, know what kinds of sequential activity really matter to what parts of the brain, and when. To this end, "template-free" or "sequence-first" approaches to detecting replay have just begun to emerge [10, 111], and are almost sure to prove disruptive.

There is also great opportunity to make sense of replay by moving beyond the hippocampus and studying what triggers replay, what biases its form and content, and what its downstream effects are. How might cells differentiate between forward and reverse replay? Does it make sense to think of some replay events as internally generated by the hippocampus, and others as externally generated by cortex? How similar are the neural circuits engaged by replay and recall? Given its sensitivity to reward, are dopaminergic systems involved in the acquisition, generation, or modification over time of replay sequences? These are just a small fraction of questions waiting to be answered outside of hippocampus.

A merger of the field—which has historically been focused on theories of memory consolidation, cognitive map formation, and spatial navigation—with the computational study of decision-making is gaining momentum. This fresh perspective may not ultimately fare any better at explaining replay, but its precisely formulated theories and relatively tight loops between theory and experiment will likely help the field to refine not only its understanding of replay, but associated concepts like planning and memory. Moreover, the focus on decision-making is likely to form a convenient bridge between the rodent literature, in which replay content is nearly always viewed as spatial (but see Stachenfeld et al. [112]), and the human literature, where abstract tasks will be preferred. The current absence of single-unit intracranial data supporting replay in humans is conspicuous and hunting for such data must surely be a top priority.

## Funding

## Disclosure Statement

None declared.

## References

1. Genzel L, et al. A consensus statement: defining terms for reactivation analysis. *Philos Trans R Soc Lond B Biol Sci*. 2020;**375**(1799):20200001.

2. Pavlides C, et al. Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *J Neurosci*. 1989;**9**(8):2907–2918.

3. van der Meer MAA, et al. Progress and issues in second-order analysis of hippocampal replay. *Philos Trans R Soc Lond B Biol Sci*. 2020;**375**(1799):20190238.

4. Tingley D, et al. On the methods for reactivation and replay analysis. *Philos Trans R Soc Lond B Biol Sci*. 2020;**375**(1799):20190231.

5. Foster DJ. Replay comes of age. *Annu Rev Neurosci*. 2017;**40**(1):581–602.

6. Levenstein D, et al. Sleep regulation of the distribution of cortical firing rates. *Curr Opin Neurobiol*. 2017;**44**:34–42.

7. de la Rocha J, et al. Correlation between neural spike trains increases with firing rate. *Nature*. 2007;**448**(7155):802–806.

8. Carey AA, et al. Reward revaluation biases hippocampal replay content away from the preferred outcome. *Nat Neurosci*. 2019;**22**(9):1450–1459.

9. Silva D, *et al*. Trajectory events across hippocampal place cells require previous experience. *Nat Neurosci*. 2015;**18**(12):1772–1779.

10. Stella F, *et al*. Hippocampal reactivation of random trajectories resembling Brownian diffusion. *Neuron*. 2019;**102**(2):450–461.e7.

11. Swanson RA, *et al*. Variable specificity of memory trace reactivation during hippocampal sharp wave ripples. *Curr Opin Behav Sci*. 2020;**32**:126–135.

12. Wilson MA, *et al*. Reactivation of hippocampal ensemble memories during sleep. *Science*. 1994;**265**(5172):676–679.

13. Kudrimoti HS, *et al*. Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. *J Neurosci*. 1999;**19**(10):4090–4101.

14. Nádasdy Z, *et al*. Replay and time compression of recurring spike sequences in the hippocampus. *J Neurosci*. 1999;**19**(21):9497–9507.

15. Skaggs WE, *et al*. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science*. 1996;**271**(5257):1870–1873.

16. Lee AK, *et al*. Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron*. 2002;**36**(6):1183–1194.

17. Buzsáki G. Hippocampal sharp waves: their origin and significance. *Brain Res*. 1986;**398**(2):242–252.

18. Buzsáki G. Hippocampal sharp wave-ripple: a cognitive biomarker for episodic memory and planning. *Hippocampus*. 2015;**25**(10):1073–1188.

19. Ji D, *et al*. Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat Neurosci*. 2007;**10**(1):100–107.

20. Pfeiffer BE, *et al*. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*. 2013;**497**(7447):74–79.

21. Karimi Abadchi J, *et al*. Spatiotemporal patterns of neocortical activity around hippocampal sharp-wave ripples. *eLife*. 2020;**9**:e51972.

22. Tingley D, *et al*. Routing of hippocampal ripples to subcortical structures via the lateral septum. *Neuron*. 2020;**105**(1):138–149.e5.

23. Klinzing JG, *et al*. Mechanisms of systems memory consolidation during sleep. *Nat Neurosci*. 2019;**22**(10):1598–1610.

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

25. Buzsáki G. Memory consolidation during sleep: a neurophysiological perspective. *J Sleep Res*. 1998;**7**(S1):17–23.

26. Foster DJ, *et al*. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*. 2006;**440**(7084):680–683.

27. Joo HR, *et al*. The hippocampal sharp wave-ripple in memory retrieval for immediate use and consolidation. *Nat Rev Neurosci*. 2018;**19**(12):744–757.

28. Rasch B, *et al*. About sleep's role in memory. *Physiol Rev*. 2013;**93**(2):681–766.

29. Csicsvari J, *et al*. Place-selective firing contributes to the reverse-order reactivation of CA1 pyramidal cells during sharp waves in open-field exploration. *Eur J Neurosci*. 2007;**26**(3):704–716.

30. Gupta AS, *et al*. Hippocampal replay is not a simple function of experience. *Neuron*. 2010;**65**(5):695–705.

31. Diba K, *et al*. Forward and reverse hippocampal place-cell sequences during ripples. *Nat Neurosci*. 2007;**10**(10):1241–1242.

32. Jackson JC, *et al*. Hippocampal sharp waves and reactivation during awake states depend on repeated sequential experience. *J Neurosci*. 2006;**26**(48):12415–12426.

33. Ólafsdóttir HF, *et al*. Hippocampal place cells construct reward related sequences through unexplored space. *Elife*. 2015;**4**:e06063.

34. Davidson TJ, *et al*. Hippocampal replay of extended experience. *Neuron*. 2009;**63**(4):497–507.

35. Ambrose RE, *et al*. Reverse replay of hippocampal place cells is uniquely modulated by changing reward. *Neuron*. 2016;**91**(5):1124–1136.

36. Singer AC, *et al*. Rewarded outcomes enhance reactivation of experience in the hippocampus. *Neuron*. 2009;**64**(6):910–921.

37. Wu CT, *et al*. Hippocampal awake replay in fear memory retrieval. *Nat Neurosci*. 2017;**20**(4):571–580.

38. O'Neill J, *et al*. Reactivation of experience-dependent cell assembly patterns in the hippocampus. *Nat Neurosci*. 2008;**11**(2):209–215.

39. Karlsson MP, *et al*. Awake replay of remote experiences in the hippocampus. *Nat Neurosci*. 2009;**12**(7):913–918.

40. Wu X, *et al*. Hippocampal replay captures the unique topological structure of a novel environment. *J Neurosci*. 2014;**34**(19):6459–6469.

41. Shin JD, *et al*. Dynamics of awake hippocampal-prefrontal replay for spatial learning and memory-guided decision making. *Neuron*. 2019;**104**(6):1110–1125.e7.

42. Xu H, *et al*. Assembly responses of hippocampal CA1 place cells predict learned behavior in goal-directed spatial tasks on the radial eight-arm maze. *Neuron*. 2019;**101**(1):119–132.e4.

43. Singer AC, *et al*. Hippocampal SWR activity predicts correct decisions during the initial learning of an alternation task. *Neuron*. 2013;**77**(6):1163–1173.

44. Dragoi G, *et al*. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature*. 2011;**469**(7330):397–401.

45. Dragoi G, *et al*. Development of schemas revealed by prior experience and NMDA receptor knock-out. *Elife*. 2013;**2**:e01326.

46. Dragoi G, *et al*. Distinct preplay of multiple novel spatial experiences in the rat. *Proc Natl Acad Sci U S A*. 2013;**110**(22):9100–9105.

47. Grosmark AD, *et al*. Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. *Science*. 2016;**351**(6280):1440–1443.

48. Liu K, *et al*. Preconfigured patterns are the primary driver of offline multi-neuronal sequence replay. *Hippocampus*. 2019;**29**(3):275–283.

49. Engel TA, *et al*. New perspectives on dimensionality and variability from large-scale cortical dynamics. *Curr Opin Neurobiol*. 2019;**58**:181–190.

50. Golub MD, *et al*. Learning by neural reassociation. *Nat Neurosci*. 2018;**21**(4):607–616.

51. Buzsáki G. *The Brain from Inside Out*. New York, NY: Oxford University Press; 2019.

52. Gridchyn I, *et al*. Assembly-specific disruption of hippocampal replay leads to selective memory deficit. *Neuron*. 2020;**106**(2):291–300.e6.

53. Wikenheiser AM, *et al*. The balance of forward and backward hippocampal sequences shifts across behavioral states. *Hippocampus*. 2013;**23**(1):22–29.

54. Cheng S, *et al*. New experiences enhance coordinated neural activity in the hippocampus. *Neuron*. 2008;**57**(2):303–313.

55. Kay K, et al. A hippocampal network for spatial coding during immobility and sleep. Nature. 2016;531(7593):185–190.

56. Oliva A, et al. Role of hippocampal CA2 region in triggering sharp-wave ripples. Neuron. 2016;91(6):1342–1355.

57. Tang W, et al. Hippocampal-prefrontal reactivation during learning is stronger in awake compared with sleep states. J Neurosci. 2017;37(49):11789–11805.

58. Mattar MG, et al. Prioritized memory access explains planning and hippocampal replay. Nat Neurosci. 2018;21(11):1609–1617.

59. Roumis DK, et al. Hippocampal sharp-wave ripples in waking and sleeping states. Curr Opin Neurobiol. 2015;35:6–12.

60. Tang W, et al. Sharp-wave ripples as a signature of hippocampal-prefrontal reactivation for memory during sleep and waking states. Neurobiol Learn Mem. 2019;160:11–20.

61. Louie K, et al. Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. Neuron. 2001;29(1):145–156.

62. Dannenberg H, et al. Modulation of hippocampal circuits by muscarinic and nicotinic receptors. Front Neural Circuits. 2017;11:102.

63. Yamamoto J, et al. Direct medial entorhinal cortex input to hippocampal CA1 is crucial for extended quiet awake replay. Neuron. 2017;96(1):217–227.e4.

64. Teles-Grilo Ruivo LM, et al. Coordinated acetylcholine release in prefrontal cortex and hippocampus is associated with arousal and reward on distinct timescales. Cell Rep. 2017;18(4):905–917.

65. Tononi G, et al. Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. Neuron. 2014;81(1):12–34.

66. Giri B, et al. Hippocampal reactivation extends for several hours following novel experience. J Neurosci. 2019;39(5):866–875.

67. Tatsuno M, et al. Methodological considerations on the use of template matching to study long-lasting memory trace replay. J Neurosci. 2006;26(42):10727–10742.

68. Leutgeb S, et al. Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. Science. 2005;309(5734):619–623.

69. Yonelinas AP, et al. A contextual binding theory of episodic memory: systems consolidation reconsidered. Nat Rev Neurosci. 2019;20(6):364–375.

70. Jadhav SP, et al. Awake hippocampal sharp-wave ripples support spatial memory. Science. 2012;336(6087):1454–1458.

71. Fernández-Ruiz A, et al. Long-duration hippocampal sharp wave ripples improve memory. Science. 2019;364(6445):1082–1086.

72. Ego-Stengel V, et al. Disruption of ripple-associated hippocampal activity during rest impairs spatial learning in the rat. Hippocampus. 2010;20(1):1–10.

73. Girardeau G, et al. Selective suppression of hippocampal ripples impairs spatial memory. Nat Neurosci. 2009;12(10):1222–1223.

74. Girardeau G, et al. Learning-induced plasticity regulates hippocampal sharp wave-ripple drive. J Neurosci. 2014;34(15):5176–5183.

75. Maingret N, et al. Hippocampo-cortical coupling mediates memory consolidation during sleep. Nat Neurosci. 2016;19(7):959–964.

76. Ciliberti D, et al. Real-time classification of experience-related ensemble spiking patterns for closed-loop applications. eLife. 2018;7:e36275.

77. Kovács KA, et al. Optogenetically blocking sharp wave ripple events in sleep does not interfere with the formation of stable spatial representation in the CA1 area of the hippocampus. PLoS One. 2016;11(10):e0164675.

78. Roux L, et al. Sharp wave ripples during learning stabilize the hippocampal spatial map. Nat Neurosci. 2017;20(6):845–853.

79. van de Ven GM, et al. Hippocampal offline reactivation consolidates recently formed cell assembly patterns during sharp wave-ripples. Neuron. 2016;92(5):968–974.

80. Packer AM, et al. Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo. Nat Methods. 2015;12(2):140–146.

81. Gauthier JL, et al. A dedicated population for reward coding in the hippocampus. Neuron. 2018;99(1):179–193.e7.

82. Schreiner T, et al. Electrophysiological signatures of memory reactivation in humans. Philos Trans R Soc Lond B Biol Sci. 2020;375(1799):20190293.

83. Liu Y, et al. Human replay spontaneously reorganizes experience. Cell. 2019;178(3):640–652.e14.

84. Constantinescu AO, et al. Organizing conceptual knowledge in humans with a gridlike code. Science. 2016;352(6292):1464–1468.

85. Quiroga RQ. Concept cells: the building blocks of declarative memory functions. Nat Rev Neurosci. 2012;13(8):587–597.

86. Staresina BP, et al. Recollection in the human hippocampal-entorhinal cell circuitry. Nat Commun. 2019;10(1):1503.

87. Kurth-Nelson Z, et al. Fast sequences of non-spatial state representations in humans. Neuron. 2016;91(1):194–204.

88. Schuck NW, et al. Sequential replay of nonspatial task states in the human hippocampus. Science. 2019;364(6447):1254–1264.

89. Ekstrom AD, et al. Cellular networks underlying human spatial navigation. Nature. 2003;425(6954):184–188.

90. Norman Y, et al. Hippocampal sharp-wave ripples linked to visual episodic recollection in humans. Science. 2019;365(6454):113–127.

91. Vaz AP, et al. Replay of cortical spiking sequences during human memory retrieval. Science. 2020;367(6482):1131–1134.

92. Isomura Y, et al. Integration and segregation of activity in entorhinal-hippocampal subregions by neocortical slow oscillations. Neuron. 2006;52(5):871–882.

93. Euston DR, et al. Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. Science. 2007;318(5853):1147–1150.

94. Wilber AA, et al. Laminar organization of encoding and memory reactivation in the parietal cortex. Neuron. 2017;95(6):1406–1419.e5.

95. Lansink CS, et al. Hippocampus leads ventral striatum in replay of place-reward information. PLoS Biol. 2009;7(8):e1000173.

96. Ólafsdóttir HF, et al. Coordinated grid and place cell replay during rest. Nat Neurosci. 2016;19(6):792–794.

97. O'Neill J, et al. Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. Science. 2017;355(6321):184–188.

98. Trimper JB, et al. Methodological caveats in the detection of coordinated replay between place cells and grid cells. Front Syst Neurosci. 2017;11:57.

99. Rothschild G, *et al*. A cortical-hippocampal-cortical loop of information processing during memory consolidation. *Nat Neurosci*. 2017;**20**(2):251–259.

100. Logothetis NK, *et al*. Hippocampal-cortical interaction during periods of subcortical silence. *Nature*. 2012;**491**(7425):547–553.

101. Jadhav SP, *et al*. Coordinated excitation and inhibition of prefrontal ensembles during awake hippocampal sharp-wave ripple events. *Neuron*. 2016;**90**(1):113–127.

102. Skelin I, *et al*. Hippocampal coupling with cortical and sub-cortical structures in the context of memory consolidation. *Neurobiol Learn Mem*. 2019;**160**:21–31.

103. Sutton RS, *et al*. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press; 2018.

104. Botvinick M, *et al*. Reinforcement learning, fast and slow. *Trends Cogn Sci*. 2019;**23**(5):408–422.

105. Mnih V, *et al*. Human-level control through deep reinforcement learning. *Nature*. 2015;**518**(7540):529–533.

106. Carr MF, *et al*. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat Neurosci*. 2011;**14**(2):147–153.

107. Johnson A, *et al*. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J Neurosci*. 2007;**27**(45):12176–12189.

108. Foster DJ, *et al*. A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus*. 2000;**10**(1):1–16.

109. Schultz W, *et al*. A neural substrate of prediction and reward. *Science*. 1997;**275**(5306):1593–1599.

110. Wood ER, *et al*. Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron*. 2000;**27**(3):623–633.

111. Maboudi K, *et al*. Uncovering temporal structure in hippocampal output patterns. *eLife*. 2018;**7**:e34467.

112. Stachenfeld KL, *et al*. The hippocampus as a predictive map. *Nat Neurosci*. 2017;**20**(11):1643–1653.