Article

Current Biology

Post-learning Hippocampal Replay Selectively Reinforces Spatial Memory for Highly Rewarded Locations

Highlights

- Rats show stronger memory for highly rewarded locations
- Disruption of hippocampal SWRs only affects memory for highly rewarded locations
- Hippocampal replay events are biased to trajectories associated with high reward

Authors

Frédéric Michon, Jyh-Jang Sun, Chae Young Kim, Davide Ciliberti, Fabian Kloosterman

Correspondence

fabian.kloosterman@nerf.be

In Brief

Michon et al. show in rats that hippocampal replay selectively enhances memory of highly rewarded locations in a familiar context. The reward-related memory enhancement is sensitive to hippocampal ripple disruption, and the proportion of replay events positively correlates with reward size and task demands.

Michon et al., 2019, Current Biology *29*, 1436–1444
May 6, 2019
https://doi.org/10.1016/j.cub.2019.03.048

CellPress

CellPress

Current Biology
Article

Post-learning Hippocampal Replay Selectively Reinforces Spatial Memory for Highly Rewarded Locations

Frédéric Michon,[1,2,3,6] Jyh-Jang Sun,[1,4,6] Chae Young Kim,[1,5,6] Davide Ciliberti,[1,2,3] and Fabian Kloosterman[1,2,3,4,7,*]
[1]NERF, Kapeldreef 75, 3001 Leuven, Belgium
[2]Brain & Cognition, KU Leuven, Tiensestraat 102, 3000 Leuven, Belgium
[3]VIB, Rijvisschestraat 120, 9052 Ghent, Belgium
[4]imec, Remisebosweg 1, 3001 Leuven, Belgium
[5]Department of Physiology, College of Medicine, Seoul National University, 28 Yeongeon-dong, 03080 Jongno-gu, Seoul, South Korea
[6]These authors contributed equally
[7]Lead Contact
*Correspondence: fabian.kloosterman@nerf.be
https://doi.org/10.1016/j.cub.2019.03.048

SUMMARY

Offline replay of hippocampal neural patterns supports the acquisition of new tasks in novel contexts, but its contribution to consolidation of salient experiences in a familiar context is unknown. Here, we show that in a highly familiar spatial memory task, large rewards selectively enhanced performance for demanding task configurations. The reward-related enhancement was sensitive to ripple-specific disruption, and the proportion of replay events positively correlated with reward size and task demands. Hippocampal replay thus selectively enhances memory of highly rewarded locations in a familiar context.

INTRODUCTION

Experiences that are salient or relevant to an organism are generally better remembered over time [1, 2]. The enhanced memory is in part due to modulation of plasticity processes, leading to stronger encoding at the time of the experience. However, selective modification and stabilization of the memory trace after initial acquisition has also been suggested to depend on the saliency of the experience [3]. Indeed, experiences associated with larger reward are followed by increased incidence of cortical markers for memory consolidation and lead to stronger and longer-lasting memory [4, 5]. These studies suggest that experienced reward modulates the prioritization of post-learning processing and consolidation of associated memory traces [6], although the neural circuit mechanisms are not well understood.

The ability to encode and remember past experiences is critically dependent on the hippocampus [7, 8]. Following initial formation, a memory trace undergoes active post-processing that is believed to reinforce hippocampal-cortical and cortico-cortical connectivity, thereby establishing a persistent trace, or schema [9]. Such system consolidation is thought to occur at a timescale of months to years in humans and at least several weeks in rodent models. The formation of a schema also leads to a reduced dependency on the hippocampus for the retrieval of remote memories, as well as the consolidation of newly formed memories associated to the pre-existing schema [10].

Population activity in the hippocampus captures the unique features of an experience [11]. Reactivation or replay of the experience-associated neuronal activity patterns in hippocampal-cortical networks during rest or sleep [12–14] is considered a prime neural substrate that enhances the acquisition and consolidation of novel information [15]. There are two lines of support for this: first, disruption of hippocampal activity specifically during replay-associated sharp-wave ripples (SWRs) impairs spatial learning [16–18]. Second, novel experiences increase replay activity, both during and after experience, and replay incidence decreases when novelty wears off [19–21]. Another factor that drives expression of replay during experience is reward consumption [22, 23]. It remains unknown, however, whether reward-driven upmodulation of hippocampal replay extends to the period after experience and supports enhanced memory consolidation of the rewarded experience.

To investigate these questions, we trained rats to learn two goal locations, one associated with a large reward and the other with a small reward, in a highly familiar setting. We show that, after a delay, rats remembered better the location associated with large reward and that the reward-related enhancement of memory was strongest for more challenging goal locations. To assess the contribution of post-learning replay activity in this paradigm, we performed ripple disruption during the delay and show that it selectively impaired memory for highly rewarded locations and, in particular, for more challenging locations. We further monitored neuronal activity in the hippocampus, and we demonstrate that post-learning replay activity of the large reward experience was enhanced.

RESULTS

To examine whether and how hippocampal replay events mediate selective reinforcement of two competing and differentially rewarded learned associations, we first developed a task in which rats were repeatedly tested for their memory of daily varying position-reward size associations (Figures 1A and S1A). Prior to the experiment, rats had been familiarized with

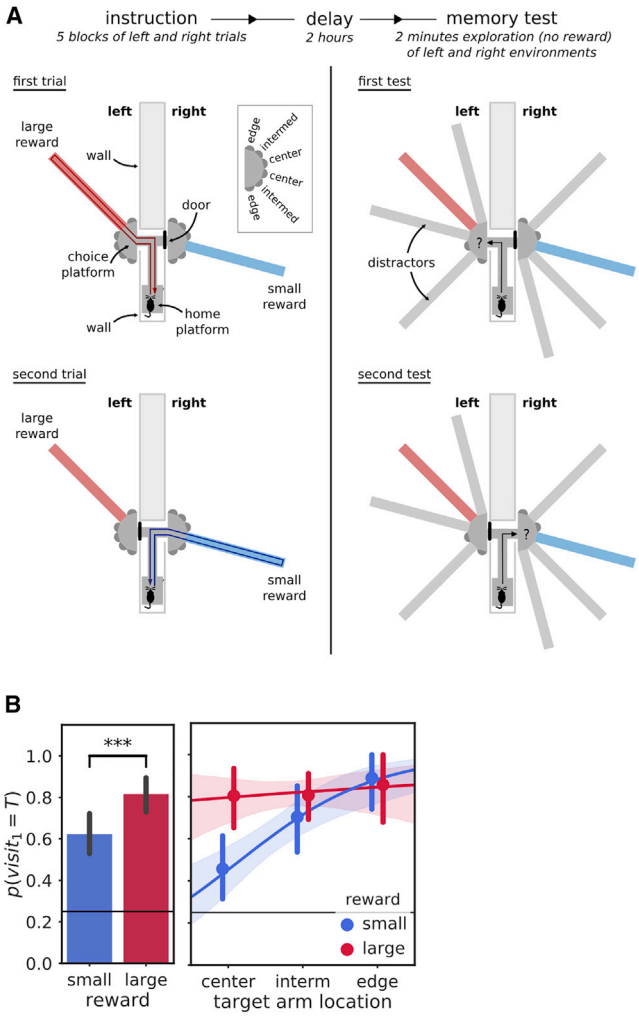

Figure 1. Dual-Environment Reward-Place Association Task

(A) The behavioral task is composed of three phases: instruction; delay; and memory probe test. During instruction, rats learn to associate small (blue) or large (red) reward amount with a specific target arm in the left and right environment. During the memory test after the delay, the preference for the target arm in presence of three distractor arms is assessed as a measure of memory. Inset: labels for target arms based on their location relative to the separating wall are shown.

(B) Left: the probability that animals first visit the target arm in the 2-min probe trial is higher in the large reward environment as compared to the small reward environment. Performance for both large and small reward environments are well above chance (black horizontal line). Right: effect of target arm location on $p(visit_1 = T)$ is shown. Filled circles represent the mean and 99% confidence interval and line and shaded region represent the mean and 95% credible region of posterior predictive samples from Bayesian generalized linear model (GLM) fit.

*** p<0.001. See also Figures S1 and S2 and Table S2.

the two environments and task rules. In training sessions, rats routinely performed instruction and reinstatement trials with a short 5 min intervening delay. During these training sessions, an equal amount of reward (3 pellets) was presented in both environments. Once the experiment started, large and small rewards were introduced and a 2-h delay was imposed between instruction and reinstatement or non-rewarded memory probe trials (see Table 1 for number of training and experimental sessions per animal). In the instruction phase of the task, rats alternately visited the left and right environment five times to retrieve either a large (9 pellets) or a small (1 pellet) reward (Figure 1A). Rats spent more time consuming the reward on the platform at the end of the target arm that was associated with a large reward as compared to small reward (mean [99% confidence interval (CI)]; large: 49.75 s [48.10, 51.50]; small: 6.44 s [5.92, 6.99]).

After the delay, rats were subjected to a memory probe test separately for the large- and small-reward environments. The animals were allowed to visit freely any of the available arms (target and three distractors) for 2 min in absence of reward. On their first journey, rats were more likely than chance level (p = 0.25) to visit the target arm ($p(visit_1 = T)$; mean [99% CI], large: 0.81 [0.73, 0.89], small: 0.62 [0.52, 0.72]; binomial test under the null hypothesis of uniform arm visit probability, large: $p = 7.1 \times 10^{-48}$, small: $p = 4.7 \times 10^{-22}$; Figure 1B). The preference for the target on the first visit was stronger for the large reward environment than the small reward environment (mean large-small difference [99% CI]: 0.19 [0.07, 0.31]; McNemar test, $H_0 : p_{small} = p_{large}$; $\chi^2 = 13.00$; p = 0.00011). A bias for the large reward environment was also reflected in a higher number of visits over the 2 min of memory probe test and in a higher average probability of visiting the target arm (Figures S1B and S1C).

The asymmetry in target arm location relative to the separating wall and the configuration of distractor arms during the memory probe test may influence the animals' performance. For example, memorizing a centrally located target arm may require a higher level of spatial integration, as compared to memorizing a target arm located at the edge for which the separating wall acts as a prominent guiding cue. On the other hand, the addition of distractor arms in the probe test may interfere with memory retrieval, especially when the distractors are positioned close to the target arm.

We first quantified the behavioral performance in the memory probe trial separately for central, intermediate, and edge target arm locations. The general reward-seeking behavior in a memory probe trial did not vary systematically with the location of the target arm (Table S2), but $p(visit_1 = T)$ and $p(T)$ did (Figures 1B and S1C). The probability of the rats to first visit the target arm was robust to the centrality of the target arm when associated with large reward but increased from central to edge target arm locations in the small reward condition. The result of logistic regression (Table S2) showed a significant positive relation between arm centrality and $p(visit_1 = T)$ only for small reward ($slope_{small}$: 1.1 [0.6, 1.7]; $slope_{large}$: 0.16 [−0.43, 0.77]), as well as significant difference in the regression slopes between large and small reward ($\Delta slope_{large-small}$: −1 [−1.7, −0.16]). The difference between large and small reward conditions was largest for central arms, whereas for edge arm locations, performance did not differ between reward conditions (Figure 1B; post hoc two proportion z test for large-small reward differences at center, intermediate, and edge arm locations, with Bonferroni corrected p value for three hypotheses; center: z = −3.73, p = 0.00058; intermediate: z = −1.35, p = 0.54; edge: z = 0.35, p = 1).

Although animals had extended experience in the maze environments (see Table S1 for the history of exposures to the maze and arms for each animal) and they were pretrained to expect the

Table 1. Overview of Training and Experimental Sessions per Animal

Animal	Training	Memory Probe	Other	Behavioral Analysis	SWR Disruption (On-Time/Delayed)	Replay Analysis
CK001	18	2	4	2		
CK002	19	2	4	2		
CK004	19	2	4	2		
CK005	14	24	28	3	11/9	
CK006	19	2	4	2		
FK010	7	7	22	7		
FK018	7	10	10	9		13
FK021	4	3	20	3		
FK022	10	6	20	6		
FK023	14	6	20	6		
FK026	13	24	36	5	9/9	
FK027	13	13	22		6/5	
FK028	9	10	10	3	5/2	4
FK030	29	2	8		1/1	
FM001	9	19	50	19		4
FM002	9	16	34	16		8
FM003	4	11	20	11		
FM004	8	13	20	13		
FM005	15	21	36	3	7/3	4
FM008	17	8	8	8		10
FM009	8	14	46	14		23
S5E001	14	16	34	6	9/1	7
S5E2	30	26	32	11	13/2	4
				$\sum = 151$	$\sum = 61/32$	$\sum = 77$

Breakdown of the number of sessions used in the behavioral analyses, sharp-wave disruption, and replay analyses. The other category includes sessions with only reinstatement trials and re-training sessions (see STAR Methods for details). See also Table S1.

same size of reward during the post-delay reinstatement trials, it is possible that motivation-related biases contribute to the pronounced difference between large and small reward conditions for the central target arm location. During the instruction phase, the run speeds toward the reward platform and the dwell time at the reward platform did not differ between central, intermediate, and edge target arm locations (Figures S2A and S2B). This suggests that animals were not more reluctant or less motivated to enter central target arms as opposed to edge target near the separating wall. To look for potential biases in the memory probe test, we analyzed which distractor arm was visited first in error trials only and did not observe a bias toward edge arms (Figure S2C).

Next, we explored the dependence of task performance on the configuration of the distractor arms relative to the target arm. Overall, $p(visit_1 = T)$ did not vary systematically either with the number of distractor arms immediately adjacent to the target arm or the compactness of distractor arms around the target, when taking into account the effect of target arm location (Figures S2D–S2F).

The observed difference in task performance between the two reward sizes may, at least in part, be mediated by memory consolidation processes during the delay phase. We hypothesized that the difference is mediated by increased reactivation of place cell sequences that represent trajectories toward the large reward. If true, we would expect that disruption of SWRs during the delay negatively affects performance in the large reward condition more strongly than in the small reward condition. To test this, we detected hippocampal SWRs online in the delay phase and used closed-loop electrical stimulation of the ventral hippocampal commissure to transiently disrupt ongoing neural activity (61 sessions in 8 animals; Table 1; Figures 2A and 2B). In control sessions (32 sessions in 8 animals; Table 1), the electrical stimulation was delayed by 100–250 ms relative to the time of SWR detection (Figure 2B). On-time stimulation of the ventral hippocampal commissure significantly reduced the ripple envelope as compared to delayed stimulation, indicating successful disruption of SWRs (Figure 2C; Mann-Whitney test: U = 1744.00, $p = 9.6 \times 10^{-12}$).

In control sessions with delayed stimulation, there was a higher probability for animals to first visit the target arm in the large reward condition as compared to the small reward condition (Figure 2D), similar to what we showed in sessions without perturbation (see Figure 1B; note that the baseline performance was lower in sessions without perturbation [i.e., Figure 1B], because of a larger fraction of early sessions in these experiments; see Figures S3A and S3B).

After SWR disruption, we found a marked decrease in $p(visit_1 = T)$ only for the large reward, but not small reward, condition. We performed binomial regression with reward and stimulation as categorical predictors (Table S3), which showed an interaction effect ($\beta_{reward \times stimulation}$: −1.9 [−3.9,

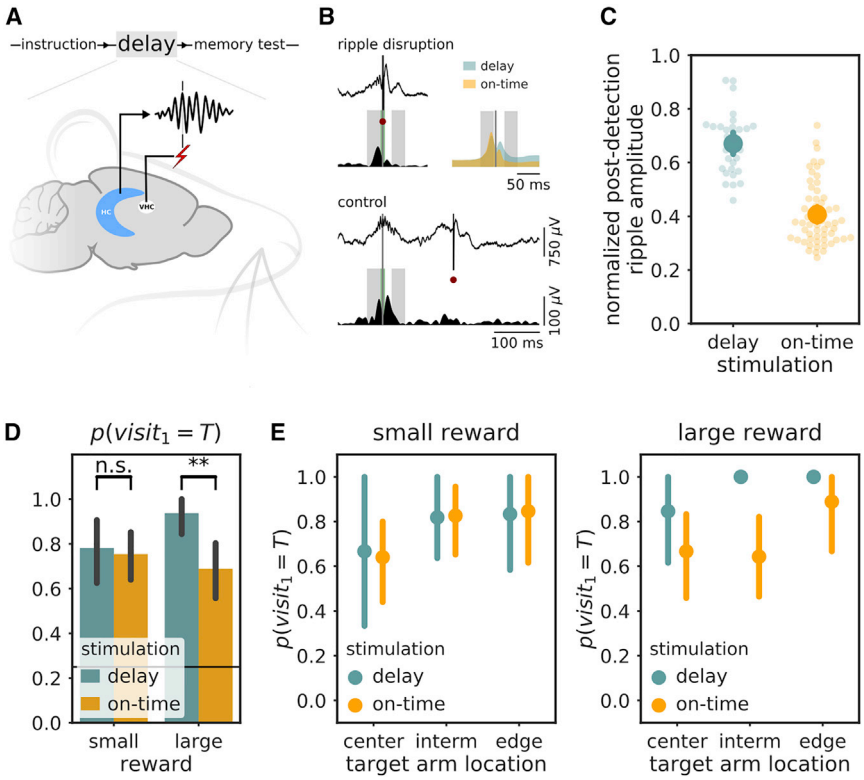

Figure 2. Disruption of SWRs during the Delay Phase Impairs Behavioral Performance

(A) Schematic of closed-loop SWR disruption approach. Ripple oscillations detected online in hippocampal area CA1 trigger electrical stimulation of the ventral hippocampal commissure (VHC) during the delay phase.

(B) Example on-time (top, left) and delayed control stimulation (bottom) from two different sessions. For both examples, the top curve represents the unfiltered field potential and the bottom plot represents the ripple oscillation envelope. Vertical gray line indicates time of SWR detection, and red dots indicate the time of VHC stimulation. Gray shaded rectangles indicate the time windows before and after detection used to assess SWR disruption. For the two sessions, the averaged ripple envelope around SWR detection is shown at top right.

(C) Quantification of ripple oscillation envelope following SWR detection for on-time and delayed control stimulation of the VHC. For each session, the average ripple envelope was computed in a 30-ms time window after detection and normalized to pre-detection envelope (time windows as shown in B).

(D) The probability of visiting the target arm on the first visit in the 2-min memory probe trial is markedly decreased after SWR disruption only for the large reward condition.

(E) Effect of SWR disruption during the delay on the probability of first visit to target depending on the location of the target arm. The strongest effect of disruption in large reward condition is observed when the target arm is located in the center or intermediate position, but not the edge position.

** p<0.01, n.s. not significant. Error bars represent 95% CI. See also Figure S3 and Table S3.

−0.12]; posterior probability that $\beta_{reward \times stimulation}$ is less than zero: p = 0.98). As a result, the difference between large and small reward conditions in control sessions ($\beta_{(large-small)|delay}$: 1.6 [−0.09, 3.4]; posterior probability that $\beta_{(large-small)|delay}$ is greater than zero: p = 0.97) was not present anymore after SWR disruption ($\beta_{(large-small)|on-time}$: −0.38 [−1.2, 0.56]). Strikingly, the decrease of $p(visit_1 = T)$ in the large reward condition after SWR disruption was most prominent when the target arm was located in the more demanding central or intermediate positions, but this effect was not present when the target arm was located at the edge (Figure 2E).

A possible mechanism to explain the sensitivity of task performance to SWR disruption specifically for the large reward condition is that a higher fraction of SWR-associated replay represents the target arm associated with large reward compared to small reward. To investigate replay expression in the task, population spiking activity was recorded in hippocampal area CA1 during instruction and delay phases (see Table 1).

Spatial representations expressed in the SWR bursts were analyzed using a neural decoding approach [24]. An encoding model was constructed that captures hippocampal spatially tuned activity during active exploration of the home and target arms in the instruction phase (see STAR Methods). Cross-validation analysis showed that the performance of decoding the animal's position in the instruction trials was high in most sessions (Figures S4A–S4D). Sessions with low decoding performance (75 percentile decoding error > 30 cm) were excluded, leaving 77 sessions in 9 animals.

First, we confirmed the presence of replay during SWR events in the instruction phase. As expected, given that animals spend more time consuming the large reward, a strong bias was found for the large reward target arm in the decoded spatial representations (Figure 3A). We next analyzed replay of spatial trajectories [14, 25] by fitting the best trajectory to the maximum-a-posteriori position estimates using weighted isotonic regression (see STAR Methods). Again, a strong bias was found for trajectory replay of the large reward target and in particular for inbound trajectories (Figures 3B–3G). Finally and consistent with a previous report [23], the bias in replay activity of the large reward target was strongest for events with positions replayed in the opposite temporal order as during behavior (reverse replay; Figure 3H).

For the remaining analyses, we focus on the delay phase and asked whether the bias for replay of the large reward target arm persisted in the delay as a possible explanation for the sensitivity to SWR disruption. During the 2-h delay, animals spent on average 85.89 min in a state of quiescence (range: 64.46–110.57 min). Within these quiet periods, the rate of SWR bursts was on average 0.25 Hz [0.22, 0.27].

We first asked whether there was a general bias in the decoded spatial representations toward the large reward target arm. For each session, we computed the mean posterior

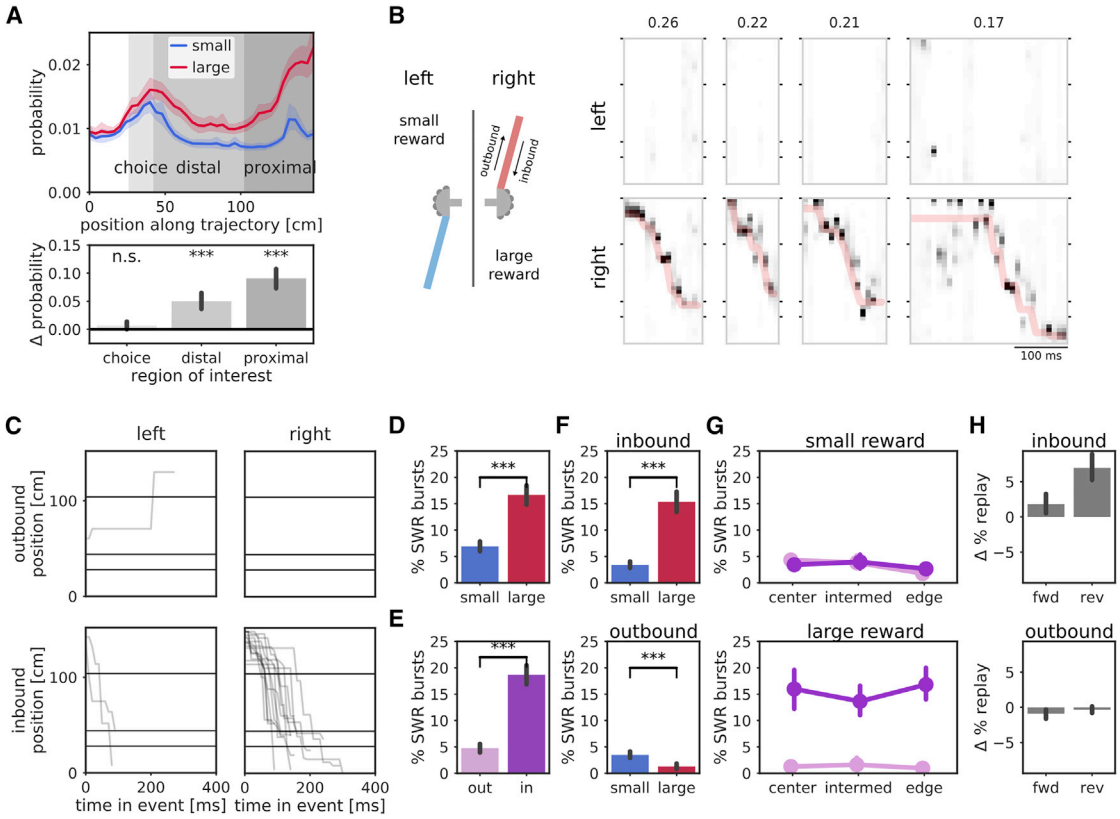

Figure 3. Hippocampal Replay Trajectories Are Biased toward Highly Rewarded Environment during Instruction

(A) Average of the per-session mean posterior probability distribution for small and large reward environments. Shaded region represents 95% CI. Bottom: quantification of the average posterior probability in three regions of interest is shown.

(B) Examples of decoded replay events in left and right environments for SWR bursts during the instruction phase in one recorded session. In this session, large reward was associated with the right environment. The shaded red line represents the fitted trajectory, and the score shown at the top corresponds to the R^2 of the isotonic regression fit weighted by the mean posterior probability at the mode.

(C) Overlay of all significant outbound and inbound trajectories in the left and right environments associated with small and large reward, respectively, from the same recording session as in (B).

(D) Mean percentage of significant replay trajectories for large and small reward environments.

(E) Mean percentage of significant inbound and outbound replay trajectories.

(F) Mean percentage of significant replay trajectories for large and small reward environments, separated by inbound and outbound direction.

(G) Mean percentage of significant replay trajectories as a function of target arm location, split by reward and inbound and outbound direction.

(H) Mean difference in percentage between large and small reward replay trajectories for forward and reverse replay events, separated by inbound (top) and outbound (bottom) trajectories. Error bar represents 99% CI.

*** p<0.001, n.s. not significant. Error bars represent 95% CI unless stated otherwise. See also Figure S4.

probability across all SWR bursts (Figure 4A, top). Across sessions, there was an overrepresentation of locations proximal to the reward site that was significantly greater for the large reward target arm as compared to the large reward target arm (Figure 4A, bottom; Kruskal-Wallis test for the large-small reward difference across proximal, distal, and choice regions of interest: H = 7.71, p = 0.021; post hoc one-sample Wilcoxon signed rank test for each region of interest with Bonferroni adjusted p values: choice Z = 1,256.00, p = 0.64, distal Z = 1,362.00, p = 1, proximal Z = 994.00, p = 0.03). These results indicate that, over all SWR bursts, there is a stronger representation of the locations close to large reward platform.

We next quantified the percentage of all SWR bursts with significant trajectory replay events in the delay phase using a weighted isotonic regression of the position estimates. Significance was tested at α = 0.01 for each event separately by comparison to the distribution of regression outcomes after 500 randomizations of the data (see STAR Methods). The average number of significant replay events was 104 ± 36 (range: 27–212), which amounted to 8.75% ± 2.52% of all SWR bursts (range: 3.2%–15.5%), which is well above the expected 1% false positive rate.

Examples of significant replay trajectories for one session are shown in Figure 4B. In this session, a majority of the significant trajectory replay events represented the environment associated with the large reward during the preceding instruction phase (Figure 4C) and, in particular, inbound trajectories (from reward site to home). Across all sessions, we observed a higher percentage of significant replay trajectories for the large reward target arm than the small reward target arm (Figure 4D; mean [99%

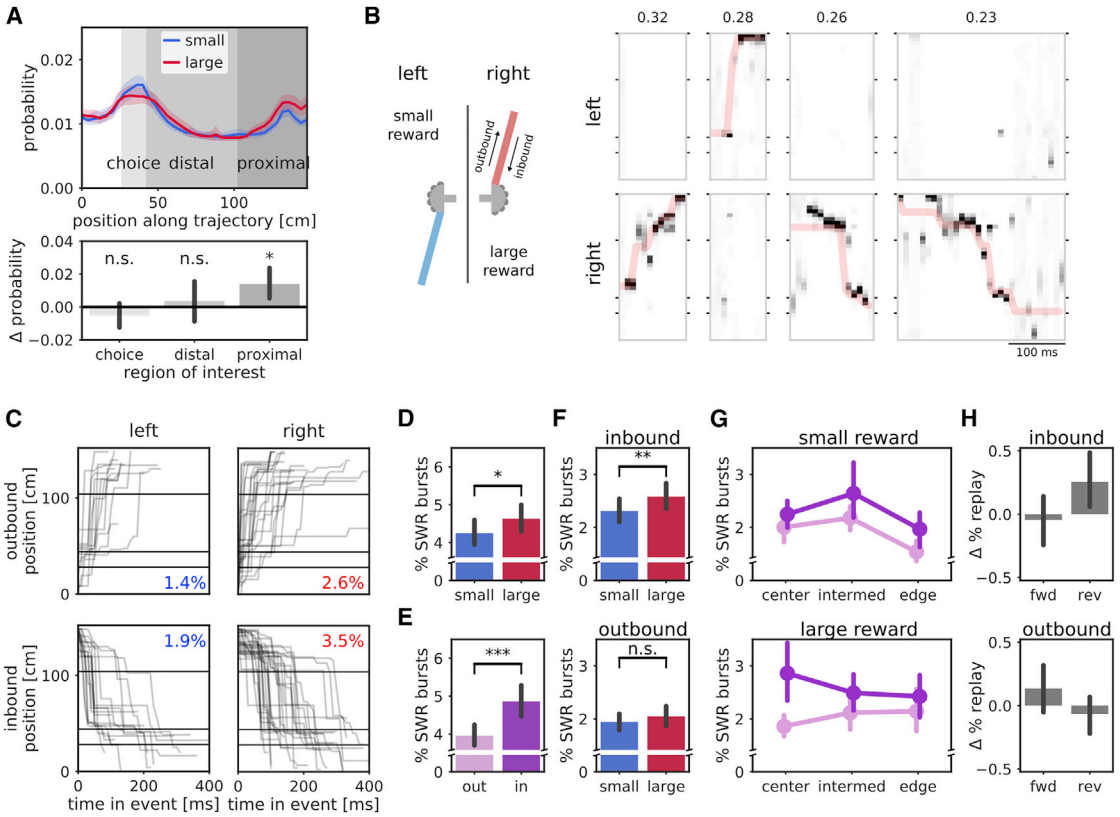

Figure 4. Hippocampal Replay during Delay Is Biased toward Environment Associated with Large Reward

(A) Average of the per-session mean posterior probability distribution for small and large reward environments. Shaded region represents 95% CI. Bottom: quantification of the average posterior probability in three regions of interest is shown.

(B) Examples of decoded replay events in left and right environments for SWR bursts during the delay phase in one recorded session. In this session, a large reward was associated with the right environment. The shaded red line represents the fitted trajectory, and the score shown at the top corresponds to the R^2 of the isotonic regression fit weighted by the mean posterior probability at the mode.

(C) Overlay of all significant outbound and inbound trajectories in the left and right environments associated with small and large reward, respectively, from the same recording session as in (B). Numbers indicate the percentage of trajectories for each category out of all candidate SWR bursts.

(D) Mean percentage of significant replay trajectories for large and small reward environments.

(E) Mean percentage of significant inbound and outbound replay trajectories.

(F) Mean percentage of significant replay trajectories for large and small reward environments, separated by inbound and outbound direction.

(G) Mean percentage of significant replay trajectories as a function of target arm location, split by reward and inbound and outbound direction.

(H) Mean difference in percentage between large and small reward replay trajectories for forward and reverse replay events, separated by inbound (top) and outbound (bottom) trajectories. Error bar represents 99% CI.

*** p<0.001, ** p<0.01, * p<0.05, n.s. not significant. Error bars represent 95% CI unless stated otherwise. See also Figure S4.

CI] percentage significant replay trajectories, small: 4.25 [3.84, 4.71], large: 4.63 [4.20, 5.10]; Wilcoxon signed-rank test: Z = 925.00, p = 0.029). Although animals spent more time at the large reward platform in the instruction phase, the percentage of replay and the difference between large and small reward during the delay phase were not related to the variation in reward platform dwell time (Figure S4E). Possible decoding biases also did not explain the bias for large reward target arm replay, as the cross-validated decoding performance did not differ between small and large reward environments (Figures S4A–S4D), the total run time, or multi-unit activity rate of the data included in the encoding model (Figures S4F and S4G).

The difference in replay between small and large reward size was largely due to a significantly higher percentage of inbound replay trajectories in the large reward target arm (Figures 4E and 4F; mean [99% CI] percentage significant inbound replay trajectories, small: mean [99% CI], 2.31 [2.04, 2.65], large: mean [99% CI], 2.58 [2.27, 2.94]; Wilcoxon signed-rank test: Z = 796.00, p = 0.0074). No significant difference between small and large reward target arms was found for outbound replay trajectories (Figure 4F; mean [99% CI] percentage significant inbound replay trajectories, small: mean [99% CI], 1.94 [1.74, 2.17], large: mean [99% CI], 2.04 [1.81, 2.30]; Wilcoxon signed-rank test: Z = 1,382.50, p = 1). Interestingly, the inbound-to-outbound replay bias for the large reward target was strongest when the target was located centrally (Figure 4G).

Finally, we decoded running direction in addition to position and determined whether each replay event during the delay phase represented forward or reverse replay. Of all significant replay trajectories, 44.95% [41.56, 48.47] had significant bias

for one of the two running directions with respect to a shuffle distribution (Monte-Carlo p value < 0.01; forward replay: 26.48% [24.40, 28.59], reverse replay: 18.51% [16.57, 20.55]). Quantification of the (large-small) reward replay difference separately for forward and reverse replay trajectories (Figure 4H) showed a bias toward reverse replay for inbound trajectories. Overall, these data indicate that there is an enrichment of inbound replay trajectories that represent the large reward target arm but use the outbound place map (e.g., reverse replay).

DISCUSSION

Altogether, our results show for the first time that post-learning hippocampal replay is modulated by previously experienced reward and contributes to the selectively enhanced consolidation of highly rewarded experiences. This effect was observed while the animals were well accustomed to both the mazes and task rules, suggesting a role for replay in the consolidation of food-place associations beyond the learning and consolidation of rules and environment representations. Our results also suggest that the replay-mediated reward-related enhancement of memory was dependent on the task difficulty. Overall, these results suggest that replay contributes to the finely tuned selective consolidation of memories under the influence of both behavioral benefit and cognitive demand.

In our behavioral paradigm, memory for small and large reward place associations is compared within the same experimental session. Although rats remembered both locations above chance after the 2-h delay, they showed better performance for the place associated with large reward. This reinforcing effect of reward on memory retention is consistent with the results of previous studies in humans [4, 5, 26, 27] and in rodents [3]. The performance boost provided by large reward in the instruction phase depended on the location of the target arm relative to the wall dividing both environments and was largest for centrally located target arms. This difference may be explained by the use of two different strategies: a non-hippocampal-dependent cue-response approach for target arms that are close to the dividing wall and a hippocampal-dependent strategy that requires a higher level of spatial integration for centrally located target arms [28] and is more challenging. If true, the large reward would then act to selectively strengthen or protect the hippocampal-dependent memory trace. This is further supported by the larger effect of hippocampal SWR disruption and the larger replay bias for more centrally located target arms. An alternative explanation for the difference in performance on central target arms for small and large rewards could be a different trade-off between the natural dislike of rats to enter open spaces (“wall hugging”) and the motivation to retrieve reward. In our experiments, rats had extensive experience with the maze apparatus and were trained repeatedly (in reinstatement trials) to expect equal size reward in the test phase after instruction. As a result, rats did not show a behavioral preference for edge arms over central arms. Thus, although the rats’ behavior in the memory probe trials may include a motivational component, we do not think it fully explains the effects of target arm location.

We found that disruption of hippocampal activity selectively during SWR bursts in rest negatively affected memory performance in the test phase. This shows that, even after extended training, when rats are highly familiar with the task rules and maze environments, hippocampal SWRs continue to contribute to learning and consolidation of the daily changing reward-place associations. Previous studies demonstrating a role of hippocampal SWRs in spatial memory consolidation using SWR-selective disruption were conducted with animals that were naive to the environment and task [17]. These studies could therefore not dissociate the contribution of hippocampal SWRs to the learning of reward-place associations from the acquisition of task rule and the familiarization to a novel spatial context.

Surprisingly, SWR disruption only affected behavioral performance in the large reward condition. No effect was observed in the small reward condition, despite the presence of hippocampal replay patterns that correspond to the small reward target arm. These results indicate that the large-reward-associated boost in memory test performance is supported by the presence of SWR bursts in the delay phase. For the less salient small reward condition, short-term (2-h) memory may be solely supported by the initial encoding in the instruction phase. Without subsequent post-processing during the delay, the small reward place associations may quickly decay after 24 h [3]. Future studies will need to reveal whether hippocampal SWRs are indeed also required for long-term (24-h) memory maintenance of the large reward place association.

Pointing to a possible mechanism for the selective vulnerability of large reward associations to hippocampal SWR disruption, we demonstrated that hippocampal replay activity during rest is biased toward the environment associated with large reward. This result is consistent with fMRI experiments in humans that revealed preferential reactivation after learning of hippocampal representations of the context associated to large reward [27]. Whether the preferred reactivation observed using fMRI is the result of spike sequence replay during short-lasting SWR bursts, however, is still an open question.

Previous studies showed that the amount of reward consumed on the maze modulates the incidence of awake replay [22], an effect that is specific to replay in the reverse temporal order compared to behavior [23]. Our results extend these findings to the rest period immediately after the learning experience, by showing a higher proportion of replay corresponding to the large reward target arm. Interestingly, we found that the higher proportion was explained by a bias for reverse replay trajectories, despite an overall higher number of forward replay events. This result differs from the equal occurrence of all preplay types after cueing (and prior to actual sampling) a rewarded goal location [29]. This suggests that preplay of unvisited cued trajectories in an otherwise known and observable environment does not share the exact same mechanisms as replay of visited trajectories that led to reward. Consistent with a previous report [30], we also found an overall spatial bias for the reward site and its vicinity. Compared to awake replay in the instruction phase, the fraction of significant trajectory replay in rest is lower. However, the absolute number of significant replay events in the 2-h delay is still an order of magnitude higher, which points to a potentially larger impact of replay events during the delay period on enhanced memory retention.

The bias toward replay for the large reward target arm in the delay period may be a consequence of stronger encoding in the hippocampal network of the rewarded experience during

the instruction phase, likely through dopamine-modulated synaptic plasticity. Indeed, it has been shown that the expression of hippocampal replay is dependent on the activation of the plasticity mediated by NMDA receptors during experience [30–32]. In addition, dopamine release varies with reward amount, and in the hippocampus, dopamine is known to promote synaptic potentiation [33, 34]. Finally, it has been shown that photostimulation of the ventral tegmental area (VTA) dopaminergic inputs to CA1 during learning enhances the post-learning reactivation of place cells with newly formed place fields [35].

Awake replay events during the instruction phase may also have contributed to the subsequent bias in replay content in rest toward the large reward target arm. Indeed, the reactivation of spike patterns during sharp wave ripples is thought to induce synaptic plasticity [36, 37]. Moreover, reverse replay rate is modulated by reward amount and reward responsive neurons in the VTA are co-activated with hippocampal reactivation during learning [38]. Consistent with this view, we found that the bias in replay for the highly rewarded arm was specific for inbound reverse trajectories initiated at the reward site both during learning and subsequent rest phase.

The demonstrated replay bias for the large reward condition in the delay period likely stabilizes the memory trace over time and hence supports higher memory retention during the test. Still, the bias is relatively modest and other factors may also be at play. Currently, we cannot exclude that qualitative differences between trajectory replay of small and large reward environments complement the quantitative differences. For example, a stronger coupling between hippocampus and VTA [27] for large reward replay events may result in a bigger change in synaptic weights. It is also conceivable that the large fraction of SWR bursts that were not classified as replay events still contribute to the enhanced memory performance for large reward-place associations. Future experiments that include replay-content-specific manipulations [39] will be able to shed light on these open questions.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Apparatus
  - Behavioral task
  - Behavioral procedure
  - Behavioral training
  - Surgical procedures
  - Electrophysiological recording
  - Online ripple detection and disruption
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavior
  - Offline ripple detection
  - Evaluation of online ripple detection and disruption
  - Candidate replay events
  - Bayesian neural decoding of hippocampal replay
  - Spatial bias in the posteriors probability distributions
  - Replay trajectory fit
  - Replay directionality
  - Statistics
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2019.03.048.

ACKNOWLEDGMENTS

We thank Jaclyn Birnbaum for help with data preprocessing and animal training, Katarzyna Bzymek for feedback on statistics, and members of the Kloosterman and Farrow labs at NERF for their comments on earlier versions of the manuscript. F.K. is funded by Research Foundation Flanders (FWO), Belgium under grant number G0D7516N and KU Leuven, Belgium C1 grant C14/17/042.

AUTHOR CONTRIBUTIONS

Conceptualization, F.M. and F.K.; Methodology, F.M., J.-J.S., C.Y.K., D.C., and F.K.; Software, D.C. and F.K.; Formal Analysis, F.M., J.-J.S., C.Y.K., and F.K.; Investigation, F.M., J.-J.S., and C.Y.K.; Writing – Original Draft, F.M., J.-J.S., C.Y.K., and F.K.; Writing – Review & Editing, F.M., J.-J.S., C.Y.K., D.C., and F.K.; Visualization, F.M., J.-J.S., and F.K.; Supervision, F.K.; Funding Acquisition, F.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 26, 2018
Revised: January 30, 2019
Accepted: March 21, 2019
Published: April 25, 2019

REFERENCES

1. Payne, J.D., Stickgold, R., Swanberg, K., and Kensinger, E.A. (2008). Sleep preferentially enhances memory for emotional components of scenes. Psychol. Sci. *19*, 781–788.
2. Wamsley, E.J., Hamilton, K., Graveline, Y., Manceor, S., and Parr, E. (2016). Test expectation enhances memory consolidation across both sleep and wake. PLoS ONE *11*, e0165141.
3. Salvetti, B., Morris, R.G.M., and Wang, S.-H. (2014). The role of rewarding and novel events in facilitating memory persistence in a separate spatial memory task. Learn. Mem. *21*, 61–72.
4. Igloi, K., Gaggioni, G., Sterpenich, V., and Schwartz, S. (2015). A nap to recap or how reward regulates hippocampal-prefrontal memory networks during daytime sleep in humans. eLife *4*, e07903.
5. Studte, S., Bridger, E., and Mecklinger, A. (2017). Sleep spindles during a nap correlate with post sleep memory performance for highly rewarded word-pairs. Brain Lang. *167*, 28–35.
6. Stickgold, R., and Walker, M.P. (2013). Sleep-dependent memory triage: evolving generalization through selective processing. Nat. Neurosci. *16*, 139–145.
7. Squire, L.R., Stark, C.E., and Clark, R.E. (2004). The medial temporal lobe. Annu. Rev. Neurosci. *27*, 279–306.
8. Zemla, R., and Basu, J. (2017). Hippocampal function in rodents. Curr. Opin. Neurobiol. *43*, 187–197.

9. Wang, S.-H., and Morris, R.G. (2010). Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. Annu. Rev. Psychol. *61*, 49–79, C1–C4.
10. Tse, D., Langston, R.F., Kakeyama, M., Bethus, I., Spooner, P.A., Wood, E.R., Witter, M.P., and Morris, R.G. (2007). Schemas and memory consolidation. Science *316*, 76–82.
11. Alme, C.B., Miao, C., Jezek, K., Treves, A., Moser, E.I., and Moser, M.-B. (2014). Place cells in the hippocampus: eleven maps for eleven rooms. Proc. Natl. Acad. Sci. USA *111*, 18428–18435.
12. Lee, A.K., and Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. Neuron *36*, 1183–1194.
13. Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature *440*, 680–683.
14. Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2009). Hippocampal replay of extended experience. Neuron *63*, 497–507.
15. Girardeau, G., and Zugaro, M. (2011). Hippocampal ripples and memory consolidation. Curr. Opin. Neurobiol. *21*, 452–459.
16. Ego-Stengel, V., and Wilson, M.A. (2010). Disruption of ripple-associated hippocampal activity during rest impairs spatial learning in the rat. Hippocampus *20*, 1–10.
17. Girardeau, G., Benchenane, K., Wiener, S.I., Buzsáki, G., and Zugaro, M.B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. Nat. Neurosci. *12*, 1222–1223.
18. Jadhav, S.P., Kemere, C., German, P.W., and Frank, L.M. (2012). Awake hippocampal sharp-wave ripples support spatial memory. Science *336*, 1454–1458.
19. 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. Nat. Neurosci. *11*, 209–215.
20. Cheng, S., and Frank, L.M. (2008). New experiences enhance coordinated neural activity in the hippocampus. Neuron *57*, 303–313.
21. Ramadan, W., Eschenko, O., and Sara, S.J. (2009). Hippocampal sharp wave/ripples during sleep for consolidation of associative memory. PLoS ONE *4*, e6697.
22. Singer, A.C., and Frank, L.M. (2009). Rewarded outcomes enhance reactivation of experience in the hippocampus. Neuron *64*, 910–921.
23. 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*, 1124–1136.
24. Kloosterman, F. (2011). Analysis of hippocampal memory replay using neural population decoding. In Neuronal Network Analysis Neuromethods, T. Fellin, and M. Halassa, eds. (Springer), pp. 259–282.
25. Karlsson, M.P., and Frank, L.M. (2009). Awake replay of remote experiences in the hippocampus. Nat. Neurosci. *12*, 913–918.
26. Nielson, K.A., and Bryant, T. (2005). The effects of non-contingent extrinsic and intrinsic rewards on memory consolidation. Neurobiol. Learn. Mem. *84*, 42–48.
27. Gruber, M.J., Ritchey, M., Wang, S.-F., Doss, M.K., and Ranganath, C. (2016). Post-learning hippocampal dynamics promote preferential retention of rewarding events. Neuron *89*, 1110–1120.
28. Ekstrom, A.D., Arnold, A.E.G.F., and Iaria, G. (2014). A critical review of the allocentric spatial representation and its neural underpinnings: toward a network-based perspective. Front. Hum. Neurosci. *8*, 803.
29. 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.
30. Dupret, D., O'Neill, J., Pleydell-Bouverie, B., and Csicsvari, J. (2010). The reorganization and reactivation of hippocampal maps predict spatial memory performance. Nat. Neurosci. *13*, 995–1002.
31. Girardeau, G., Cei, A., and Zugaro, M. (2014). Learning-induced plasticity regulates hippocampal sharp wave-ripple drive. J. Neurosci. *34*, 5176–5183.
32. Silva, D., Feng, T., and Foster, D.J. (2015). Trajectory events across hippocampal place cells require previous experience. Nat. Neurosci. *18*, 1772–1779.
33. Lak, A., Stauffer, W.R., and Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. Proc. Natl. Acad. Sci. USA *111*, 2343–2348.
34. Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. Science *307*, 1642–1645.
35. McNamara, C.G., Tejero-Cantero, Á., Trouche, S., Campo-Urriza, N., and Dupret, D. (2014). Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. Nat. Neurosci. *17*, 1658–1660.
36. Sadowski, J.H., Jones, M.W., and Mellor, J.R. (2016). Sharp-wave ripples orchestrate the induction of synaptic plasticity during reactivation of place cell firing patterns in the hippocampus. Cell Rep. *14*, 1916–1929.
37. Roux, L., Hu, B., Eichler, R., Stark, E., and Buzsáki, G. (2017). Sharp wave ripples during learning stabilize the hippocampal spatial map. Nat. Neurosci. *20*, 845–853.
38. Gomperts, S.N., Kloosterman, F., and Wilson, M.A. (2015). VTA neurons coordinate with the hippocampal reactivation of spatial experience. eLife *4*, e05360.
39. Ciliberti, D., Michon, F., and Kloosterman, F. (2018). Real-time classification of experience-related ensemble spiking patterns for closed-loop applications. eLife *7*, e36275.
40. Kloosterman, F., Davidson, T.J., Gomperts, S.N., Layton, S.P., Hale, G., Nguyen, D.P., and Wilson, M.A. (2009). Micro-drive array for chronic in vivo recording: drive fabrication. J. Vis. Exp. e1094.
41. Nguyen, D.P., Layton, S.P., Hale, G., Gomperts, S.N., Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2009). Micro-drive array for chronic in vivo recording: tetrode assembly. J. Vis. Exp. e1098.
42. Ciliberti, D., and Kloosterman, F. (2017). Falcon: a highly flexible open-source software for closed-loop neuroscience. J. Neural Eng. *14*, 045004.
43. Millman, K.J., and Aivazis, M. (2011). Python for scientists and engineers. Comput. Sci. Eng. *13*, 9–12.
44. Kloosterman, F., Layton, S.P., Chen, Z., and Wilson, M.A. (2014). Bayesian decoding using unsorted spikes in the rat hippocampus. J. Neurophysiol. *111*, 217–227.
45. Sodkomkham, D., Ciliberti, D., Wilson, M.A., Fukui, K., Moriyama, K., Numao, M., and Kloosterman, F. (2016). Kernel density compression for real-time Bayesian encoding/decoding of unsorted hippocampal spikes. Knowl. Base. Syst. *94*, 1–12.
46. Salvatier, J., Wiecki, T.V., and Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. PeerJ Comput. Sci. *2*, e55.
47. Oliphant, T.E. (2007). Python for scientific computing. Comput. Sci. Eng. *9*, 10–20.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Rat: Long Evans	Janvier, France	Cat# RjOrl:LE
Software and Algorithms		
Scientific Python	[43, 47]	http://www.scipy.org
Falcon real-time neural signal processing software	[42]	http://www.bitbucket.org/kloostermannerflab
Neural decoding without spike sorting	[44, 45]	N/A
Other		
Digilynx SX acquisition system with HS-36 analog headstage and Cheetah software	Neuralynx	http://www.neuralynx.com
12 μm polyimide-insulated nickel-chrome tetrode wire	Sandvik	Cat# NI055820
60 μm polyimide-insulated stainless steel stimulation wire	California Fine Wire	http://www.calfinewire.com
Stimulus generator	MultiChannel Systems	STG4000 Series

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Fabian Kloosterman (fabian.kloosterman@nerf.be).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A total of 23 male Long Evans rats food deprived to 85%–90% of the free-feeding weight were used in this study. At the start of behavioral training procedures, animals were 7-10 weeks old. Thirteen animals received an implant for electrical recording, and 10 rats did not undergo surgical procedures and were only tested behaviorally. Table 1 provides an overview of the experimental sessions for each animal.

All experiments were carried out in accordance with protocols approved by KU Leuven animal ethics committee (P119/2015) in accordance with the European Council Directive, 2010/63/EU. Animals in experiments were housed separately in individually ventilated cages (IVC) with *ad libitum* access to water and controlled intake of standard food pellets. Health status and body weight are checked daily by the experimenters and dedicated animal care personnel.

METHOD DETAILS

Apparatus

The behavioral testing apparatus was located in one of two 4x4 m rooms with black walls and distinctive, orienting cues on each of the walls. The apparatus was elevated 40 cm off the ground and consisted of a home platform that via a short 30 cm track gave access to the left and right side of the room (Figure 1A). The left and right environments were separated by 120 cm high dividers. In each environment, a choice platform gave access to at most 6 radially emanating 90 cm long arms with neighboring arms separated by 30°. Access from the home platform to the two environments was controlled by a door that was manually positioned by the experimenter.

Behavioral task

The goal of the task was for the rat to associate one of the 6 arms in each environment with reward. In each daily session, one environment was associated with large reward (9 pellets) and the other with small reward (1 pellet). After an acquisition phase, in which the animal could explore the rewarded target arms across 5 instruction trials per environment, and after a subsequent 2h delay, the rats were tested for their memory of the daily reward-location association in the presence of three additional distractor arms (Figure 1A). Across sessions, the location of the target and the large/small reward assignment of the target arm were varied pseudo-randomly (Figure S1). The location of the distractor arms were chosen pseudo-randomly, such that at least one distractor arm was directly adjacent to the target arm and the target arm of the previous session was not part of the distractors.

Behavioral procedure

In the instruction phase, only the two target arms were physically present in the two environments. The instruction was given in 5 blocks of 2 trials, one trial for the left and right environment. In each trial, the rat started at the home platform and was given exclusive access to one of the environments. A new trial started as soon as the rat consumed the reward at the end of the target arm and returned to the home platform. The presentation order of the environments within the blocks was constant within a session and randomized across sessions.

Following the instruction phase and before the test phase, the rat was removed from the maze apparatus for either a short (at most 15 minute) delay during training or a 2h delay in experimental sessions. During the delay, the rat was either returned to its home cage (when no electrophysiological recordings were performed) or placed in a 40x40 cm rest box with 60 cm high wall that was located inside the room.

After the delay, rats were subjected to three reinstatement trials separately for the two environments in the presence of distractor arms. In each reinstatement trial, rats were rewarded for visiting the target arm with equal amount (3 pellets), regardless of the reward magnitude associated with the target arm in the instruction phase. The reinstatement trials served to train the rats to search for reward in the target and to ignore non-rewarded distractor arms. In approximately 50% of sessions, the first reinstatement trial was replaced by a two-minute long unrewarded memory probe trial for each of the two environments (for number of sessions per animal, see Table 1). This procedure was followed to make sure that the rats retained their motivation to search for reward in the memory probe trials.

Behavioral training

Prior to experimental sessions, animals were gently handled by the experimenter at least 5 minutes per day for 7 days and pre-trained to run back-and-forth on an elevated linear track (40 cm high and 90 cm long) for food rewards (3 pellets). Pre-training continued until the rats executed at least 20 laps within 10 minutes for three sessions in a row. The rats were then familiarized with the overall structure of the reward-place association task. During these daily training sessions, both environments were associated with an equal amount of reward (3 pellets) in the instruction phase, the delay was kept short (5-15 minutes), and rats were subjected to three reinstatement trials after the delay. The familiarization lasted until the rats went straight to the target arm on the first reinstatement trial in both environments for at least 3 sessions in a row (for number of training sessions, see Table 1). Animals that did not reach criterion were excluded.

Surgical procedures

Custom-designed 3D-printed micro-drive array [40, 41], carrying up to 24 tetrodes (four twisted 0.012 mm polyimide-insulated nickel-chrome wires, blunt cut; Sandvik, Kista, Sweden) and 3 stimulation electrodes (two twisted 0.06 mm polyimide-insulated stainless steel wires, blunt or angled cut; California Fine Wire, Grover Beach, CA), was surgically attached to the rat skull using standard aseptic techniques. After initial induction of anesthesia (5% isoflurane in induction chamber), the rat's scalp was shaved and the rat's head was securely mounted in a stereotaxic frame. Throughout the surgical procedure, anesthesia was maintained using 0.5%–2% isoflurane administered through a nose cap. The level of anesthesia was kept constant by continuously monitoring the breathing and heart rate and the concentration of isoflurane was adjusted when needed. The body temperature of the rat was measured using a rectal probe and kept constant using an electrical heating pad. Following disinfection of the scalp with iodine and isopropyl alcohol, the skull was exposed and 8-10 anchoring bone screws (Fine Science Tools, Heidelberg, Germany) were inserted around the perimeter. One of the bone screws was used as a ground for electrophysiological recordings. Next, craniotomies were made to provide access to the hippocampus for recording electrodes (center coordinates: 4 mm posterior to Bregma, 2.5 mm right from the midline) and to the ventral hippocampal commissure for stimulation electrodes (center coordinates: 1.3 mm posterior to Bregma, 0.9 mm right from the midline). After removal of the dura, the micro-drive array was fixed in place to the anchoring screws using light-curable dental cement (SDI, Bayswater, Australia).

Electrophysiological recording

After 1 week of post-operative recovery, stimulation and recording electrodes were lowered into position over the course of 2-3 weeks. Recording electrodes were moved into the hippocampal area CA1 pyramidal cell layer. Wide-band (0.1 Hz - 6 kHz) signals were sampled at 32 kHz and digitized using a 128-channel data acquisition system (Digilynx SX acquisition system with HS-36 analog headstage and Cheetah software; Neuralynx, Bozeman, MO). Waveform snippets of online detected spike events in the band-pass filtered signal (600 Hz - 6 kHz) were saved to hard disk at the original sampling rate. The position and head direction of the animal were tracked and captured at 50 Hz using an overhead video camera and colored LEDs mounted on the headstage. Implanted animals were retrained in the task without delay during the time electrodes were lowered in place, and they only entered the experiment after reaching criterion (see Behavioral training).

Online ripple detection and disruption

A network stream of digitized multi-channel samples from the DigiLynx acquisition system was fed into a quad-core workstation that supports hardware-based multithreading. The workstation ran the open source real-time processing software Falcon [42] and generated TTL pulses via an isolated Digital Input/Output (DIO) module (USB-4750, Advantech Benelux, Breda, the Netherlands) or Arduino UNO.

Falcon software was configured with a data processing graph that detected transient increases in hippocampal ripple power in the incoming data stream. Raw signals from 1-3 electrodes were filtered between 135 and 255 Hz using a Chebyshev type-II IIR filter (order = 20) that was designed to minimize group delay (minimum delay of 6.5 ms at 200 Hz). Ripple power was summed across electrodes and ripple events were detected when the summed power exceeded $f \times mad(t) + \mu(t)$, where $\mu(t)$ and $mad(t)$ are running estimates of the mean and mean absolute deviation calculated on the ripple power, and f is a multiplier set to a value in the range [5, 13]. Estimates of $\mu(t)$ and $mad(t)$ were computed using an exponential moving average filter with span set to a value in the range of [5, 15] seconds. Estimates were not updated during a 50 ms window after each detection.

To reduce spurious detections induced by movement artifacts or stimulus-evoked responses, the same ripple detection procedure was applied to signals from 1-2 electrodes that were positioned in the cortex overlying the hippocampus. Hippocampal ripple detections were ignored if they fell within a time window (−40 ms to 1.5 ms) around spurious detections in the cortical signals.

Detection of a ripple event triggered a TTL pulse on the DIO card or Arduino UNO that was sent to the Diglynx system for logging and to a constant-current stimulator (MultiChannel System, Reutlingen, Germany) for electrical stimulation of the ventral hippocampal commissure. Amplitude of the biphasic electrical pulses (0.2 ms duration) varied from 100-500 μA and was set in each session to the lowest amplitude that resulted in consistent disruption of hippocampal ripple events. A lock-out period was used to limit the output stimulation frequency to 2 Hz and avoid overstimulation. In control experiments, electrical stimulation was delayed by 100-250 ms relative to the detection of a ripple event. Detections that occurred during the delay period were discarded. No systematic change in the evoked response was observed over the 2 hours of stimulation (Figure S3H), indicating that the SWR disruption procedure did not induce long-term plasticity.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of neural and behavioral data was performed using Python and its scientific extension modules [43], augmented with custom Python and C++ toolboxes.

Behavior

In the instruction trial, the average running speed to and from the reward platform was quantified only for the implanted animals, based on video tracking of the head-mounted LEDs. Average speed was computed over the journey that started when the animal left home and ended when the animal reached the reward platform at the end of a target arm (and vice-versa for the homebound journey).

In the memory probe trial, the number and pattern of visits to the target and distractor arms were quantified as measures of performance in the reward-place associative task. A visit to an arm was only counted if the animal reached the reward platform at the end of an arm. We defined the following quantities and (conditional) probabilities to characterize the reward-seeking behavior in the probe trial: $\boldsymbol{N_{visits}}$: the total number of arm visits in the 2 minute memory probe trial. $\boldsymbol{p}(\boldsymbol{visit}_1 = \boldsymbol{T})$: the across session mean probability that the first visit is on target. $\boldsymbol{p}(\boldsymbol{T})$: the across session mean probability that a visit is made to the target, computed by averaging the equivalent per session $p(T)$.

Offline ripple detection

The local field potential recorded from 1-3 tetrodes was downsampled from 32 kHz to 4 kHz and filtered in the ripple frequency band (140-225 Hz). The ripple envelope was computed as the absolute value of the Hilbert-transformed ripple signal, averaged across the recording sites and smoothed with a Gaussian kernel (bandwidth 15 ms). Slow trends in the ripple envelope were removed using a moving median filter (window length 3 s). Finally, start and end times of ripple events were detected when the detrended ripple envelope exceeded a low threshold of $\mu + 0.5 \times \sigma$ and the maximum envelope exceeded a high threshold of $\mu + 8 \times \sigma$. Here, μ and σ represent the mean and standard deviation of the detrended envelope. Ripple events that were separated by less than 20 ms were merged into one, and events with a duration shorter than 40 ms were excluded.

Evaluation of online ripple detection and disruption

Online SWR detection accuracy was verified in a group of sessions in which SWRs were detected during the delay, but no on-time or delayed electrical stimulation was performed. Online detected SWRs were compared to offline defined SWRs to identify the fraction of offline SWRs that were correctly detected online (true positive rate or TPR) and the fraction of online detections that did not correspond to an offline defined SWR (false discovery rate, FDR) (Figure S3C). The online SWR detection rate varies across sessions due to differences in signal-to-noise ratio and experimenter-set detection threshold. Higher detection rates are accompanied by increasing true positive and false discovery rates. The true positive rate saturated around 1 Hz detection rate and higher detection rates were accompanied by an increasing false discovery rate (Figure S3D). A maximum detection rate of 1 Hz was set as a selection criterion for the sessions with on-time or delayed stimulation, in which the stimulus artifact interfered with proper characterization of online detection performance.

To assess the disruption of SWRs after on-time closed-loop stimulation of the VHC, we adopted the approach of [16]. After removal of the stimulation artifact and cubic interpolation of the signal in a 10 ms window around the time of stimulation, the ripple envelope was computed as before. For delayed stimulation control sessions, signal interpolation was performed for a 10 ms time window around both the detection and stimulation times. The mean ripple envelope after each detection/stimulation was

computed in a 30 ms time window (from +20 ms to +50 ms) and normalized to the mean ripple envelope in a 30 ms time window (from −35 ms to −5 ms) prior to detection/stimulation.

Candidate replay events

A smoothed multi-unit activity (mua) rate histogram (5 ms bin size, Gaussian kernel with 15 ms bandwidth) was computed from all unsorted spikes recorded from the hippocampus with a peak amplitude larger than 60 μA. The rate was detrended using a moving median filter (window length 3 s). Transient bursts in the detrended mua were defined using a double threshold procedure, where the upper threshold $\mu + 4\times\sigma$ determines if a burst occurred and the lower threshold $\mu + 0.5\times\sigma$ determines burst start and end time. Bursts that were separated by less than 20 ms were merged into one, and bursts with a duration shorter than 80 ms were excluded. Candidate replay events were defined as mua bursts that overlapped with a ripple event and which occurred while the animal was immobile (speed < 5 cm/s).

Bayesian neural decoding of hippocampal replay

A Bayesian neural decoding approach was used to analyze the spatial information content of candidate replay events [14, 24]. Per recording session, an encoding model that relates hippocampal spiking activity to the animal's position in the home and rewarded target arms was constructed from the data acquired in the instruction phase of the task. Only spikes emitted during run epochs (speed > 10 cm/s) with a minimum spike amplitude of 60 μV were incorporated into the model. Tetrodes with a mean spiking rate during run epoch below 0.1 Hz were excluded.

We used a decoding approach that directly relates spike amplitude features to position without prior spike sorting [44]. Under the assumption that all spikes on a tetrode occur conditionally independent of past spikes and that the firing rate is determined by position in the maze, the hippocampal activity on a single tetrode can be modeled as a marked temporal Poisson process that is fully characterized by the rate function $\lambda(\mathbf{a},x)$, where $\mathbf{a}$ represents the vector of spike amplitudes and x represent position in the maze. The likelihood of observing a set of spikes with an amplitude of $\mathbf{a}_{1:n}$ in time interval Δ for a given position is then expressed as [44]:

$$P(\mathbf{a}_{1:n}\,|\,x) = \Delta^n \left[\prod_{i=1}^{n} \lambda(\mathbf{a}_i, x)\right] e^{\Delta\lambda(x)} \tag{1}$$

And the joint likelihood across K tetrodes is obtained by product of the single tetrode likelihoods: $P(\mathbf{a}^{1:K}\,|\,x) = \prod_{k=1}^{K} P(\mathbf{a}_{1:n_k}\,|\,x)$.

We used a compressed kernel density estimator to evaluate the rate function $\lambda(\mathbf{a},x)$ and the marginal rate function $\lambda(x)$ from their component spike count and position occupancy probability distributions [45]. The bandwidth of the Gaussian kernel was set to 30 μV for spike amplitude and 5 cm for position. Mahalanobis distance threshold for compression was set to 1.0 for an acceptable trade-off between decoding accuracy and computation time.

To perform neural decoding for the spiking activity recorded on K tetrodes in time window Δ and estimate the posterior probability distribution over position (sampled at a regular grid with 4 cm spacing) from the likelihood we resort to Bayes' rule:

$$P(x\,|\,\mathbf{a}^{1:K}) = \frac{P(\mathbf{a}^{1:K}\,|\,x)P(x)}{P(\mathbf{a}^{1:K})} \tag{2}$$

where a uniform prior $P(x)$ is used.

To evaluate the performance of the decoder, we used a five-fold cross-validation procedure in which four out of five instruction trial blocks were used to build the encoding model and the remaining instruction trial block was used for decoding the animal's position on the maze in Δ = 100 ms time bins. The decoding error was defined as the distance along the maze between the estimated and real position. For each session, we evaluated the decoding error distribution separately for left and right environments, as well as for the complete maze (Figures S4A–S4D). Only sessions with a good decoding performance during the instruction phase (75 percentile of decoding error distribution is below 30 cm) were selected for subsequent analysis of hippocampal replay (Figures S4A–S4D).

Spatial bias in the posteriors probability distributions

To assess spatial bias in the decoded spatial representation across all candidate events, the mean of the posterior probabilities was computed separately for the small and large reward environments. For statistical comparison, mean posterior probability for three regions of interest were computed: proximal target arm (within 50 cm of the reward site), distal target arm (the remaining 60 cm of the target arm) and the choice platform.

Replay trajectory fit

Candidate replay events were split into Δ = 10 ms time bins and spiking activity (amplitude > 60 μV) in each bin was used to perform decoding as described above. Next, separately for left and right target arms, weighted isotonic regression was performed on the maximum-a-posteriori (MAP) position estimates with posterior probabilities as weights. A goodness-of-fit score was defined that combined both the posterior probabilities and the R^2 of the regression: $score = 1/T\sum_{t=1}^{T} P_{MAP,t} \times R^2$, where T is the number of time bins in the event, and $P_{MAP,t}$ is the posterior probability associated with the MAP estimate in time bin t. The regression was performed twice to fit both a monotonically increasing and a monotonically decreasing trajectory to the MAP estimates, and only the best fitting trajectory with the highest score was retained.

For each event, the goodness-of-fit score is compared to the distribution of scores constructed from 500 pseudo-random events in which each posterior was randomly drawn from the complete set of candidate replay events [14]. Only candidate replay events with a Monte-Carlo p value < 0.01 were considered to contain significant trajectory replay.

Replay directionality

For each significant replay trajectory, the run direction was decoded using the same Bayesian neural decoding approach as used for position. Cross-validation showed that decoding performance for run direction was high (Figure S4D). For each replay event, a direction bias was computed separately for the small and large reward environment as the mean difference in posterior probability between the inbound and outbound run directions across time bins. The direction bias is compared to the distribution of biases obtained from 1000 pseudo-random events in which each posterior is randomly assigned a direction. Replay events with a significant direction bias (Monte-Carlo p value < 0.01) were classified as either 'forward' or 'reverse', depending on whether the decoded direction matched the direction of the replay trajectory.

Statistics

To test a difference in means between two paired samples, we used the Wilcoxon signed rank test. To test for a difference in two sample proportions, we used either the McNemar test (for paired samples) or the two proportion z-test.

To analyze the dependence of behavioral metrics on predictor variables, we fitted Bayesian generalized linear models (GLMs) using the PyMC3 package for Python [46]. We applied a Poisson regression model (with log link function) for the number of arm visits, a logistic regression model (with logit link function) for $p(visit_1 = T)$ and an ordinary linear regression model for $p(T)$.

We used the following model to study the relation between target arm location (centrality), reward and behavior in the memory probe trial:

$$E(y) = g^{-1}(\beta_0 + \beta_1 \times reward + \beta_2 \times centrality + \beta_3 \times reward \times centrality)$$

With the equivalent per-reward model equations:

$$E(y_{small}) = g^{-1}(\beta_0 + \beta_2 \times centrality) = g^{-1}(intercept_{small} + slope_{small} \times centrality)$$

$$E(y_{large}) = g^{-1}(\beta_0 + \beta_1 + (\beta_2 + \beta_3) \times centrality) = g^{-1}(intercept_{large} + slope_{large} \times centrality)$$

Here, β_c are the model parameters, g^{-1} is the link function. The categorical reward predictor variable was coded as small = 0 and large = 1. The centrality predictor variable was treated as an ordinal variable with [center, intermediate, edge] coded as [0, 1, 2].

We used the following model to study the relation between on-time and delayed SWR disruption, reward and behavior in the memory probe trial:

$$E(y) = g^{-1}(\beta_0 + \beta_1 \times reward + \beta_2 \times disruption + \beta_3 \times reward \times disruption)$$

Here, β_c are the model parameters, g^{-1} is the link function. The categorical reward and disruption predictors were dummy coded (reward: small = 0 and large = 1; disruption: delayed = 0, on-time = 1).

Model fitting and inference were performed using Markov chain Monte Carlo (MCMC) sampling methods in PyMC3 (specifically, the No-U-Turn Sampler). Broad normal distributions were used as priors on the parameters. A summary of the model fitting results is presented in Tables S2 and S3, where we list for the relevant (untransformed) model parameters: the posterior mode, the 95% highest posterior density (95% HPD) interval, the posterior probability that the parameter is less or greater than zero and the Gelman-Rubin diagnostic R that indicates good convergence of the Markov chain when its value is close to 1.

DATA AND SOFTWARE AVAILABILITY

Falcon software for closed-loop SWR detection is publicly available at http://www.bitbucket.org/kloostermannerflab. The data and analysis routines used in this study are available on request, please contact the corresponding author.