nature neuroscience

SUPPLEMENTARY INFORMATION

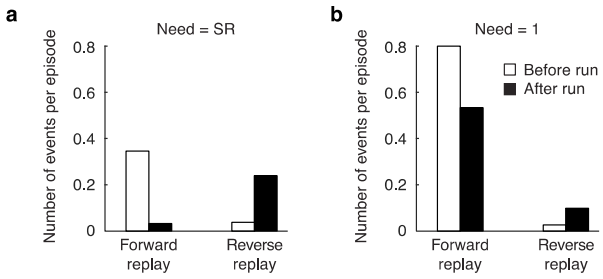
https://doi.org/10.1038/s41593-018-0232-z

In the format provided by the authors and unedited.

Prioritized memory access explains planning and hippocampal replay

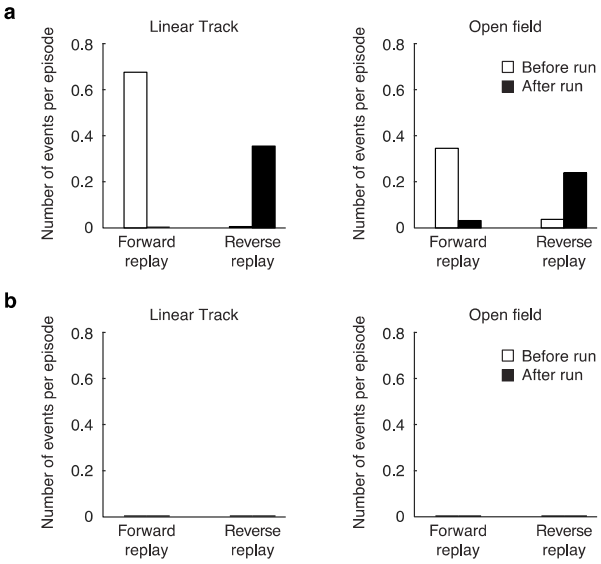
Marcelo G. Mattar[1]* and Nathaniel D. Daw[1,2]

[1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA. [2]Department of Psychology, Princeton University, Princeton, NJ, USA.
*e-mail: mmattar@princeton.edu



Supplementary Figure 1

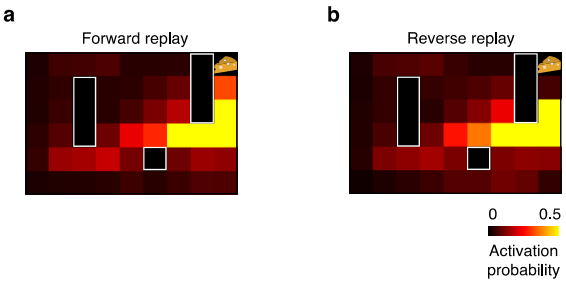
The need term is required for channeling backups along depth-first trajectories.
(a) Number of significant forward and reverse sequences on the open field when setting the need term to the Successor Representation (as predicted by the theory). *(b)* Number of significant forward and reverse sequences on the open field when setting the need term to be one at every state. Notice the reduction of reverse sequences at the end of a run due to a breadth-first pattern of value propagation.



Supplementary Figure 2

A positive need term is required for sequential value propagation.

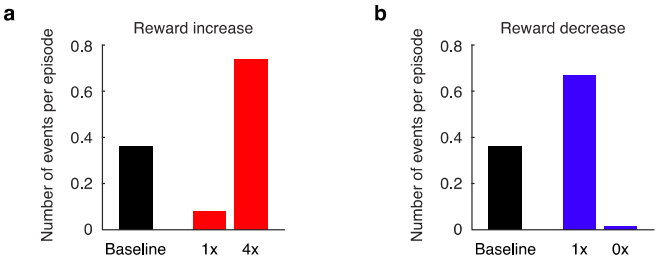
(a) Number of significant forward and reverse sequences on linear track (*left*) and on open field (*right*) when setting the need term to the Successor Representation (as predicted by the theory). *(b)* Number of significant forward and reverse sequences on linear track (*left*) and on open field (*right*) when setting the need term to be one at the agent's location and zero elsewhere. Notice the complete absence of both forward and reverse sequences due repeated replay of actions on the agent's location.



Supplementary Figure 3

Replay over-represents reward locations in both reverse and forward sequences.

(a) Probability that a state is part of a forward sequence at a given episode. Notice that because of the random starting location, the observed results cannot be explained by initiation bias. Notice also that locations corresponding to the final turn toward the reward were emphasized even more than locations nearer the reward itself, a consequence of the gain term being higher where there is a greater effect on behavior. *(b)* Probability that a state is part of a reverse sequence at a given episode. Notice that because of the random starting location, the observed results cannot be explained by initiation bias. Reverse replay tends to concentrate near the reward. Notice that the higher activation probability for reverse events is due to a combination of a reward-location bias and initiation bias, given that reverse sequences tend to start near the reward, where the agent usually is.



Supplementary Figure 4

Differential modulation of reverse replay in increased/decreased reward conditions.
(a) Increased reward condition where the reward encountered by the agent was four times larger in half of the episodes. In comparison to the baseline (1x-1x) reward setting (*left*), there was a lower number of significant reverse sequences for 1x rewards (*middle*) and a greater number of significant reverse sequences for 4x rewards (*right*). The results presented in Fig. 5c correspond to a ratio between the two red bars (4x/1x). *(b)* Decreased reward condition where the reward encountered by the agent was zero in half of the episodes. In comparison to the baseline (1x-1x) reward setting (*left*), there was a greater number of significant reverse sequences for 1x rewards (*middle*) and a lower number of significant reverse sequences for 0x rewards (*right*). The results presented in Fig. 5e correspond to a ratio between the two blue bars (0x/1x).