

# Systems consolidation impairs behavioral flexibility

Sankirithana Sathiyakumar,<sup>1,2</sup> Sofia Skromne Carrasco,<sup>1</sup> Lydia Saad,<sup>1</sup>  
and Blake A. Richards<sup>1,2,3,4,5</sup>

<sup>1</sup>Department of Biological Sciences, University of Toronto Scarborough, Toronto, Ontario M1C 1A4, Canada; <sup>2</sup>Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario M5S 3G5, Canada; <sup>3</sup>Mila, Montréal, Quebec H2S 3H1, Canada; <sup>4</sup>Department of Neurology and Neurosurgery, McGill University, Montréal, Quebec H3A 2B4, Canada; <sup>5</sup>School of Computer Science, McGill University, Montréal, Quebec H3A 2A7, Canada

Behavioral flexibility is important in a changing environment. Previous research suggests that systems consolidation, a long-term poststorage process that alters memory traces, may reduce behavioral flexibility. However, exactly how systems consolidation affects flexibility is unknown. Here, we tested how systems consolidation affects: (1) flexibility in response to value changes and (2) flexibility in response to changes in the optimal sequence of actions. Mice were trained to obtain food rewards in a Y-maze by switching nose pokes between three arms. During initial training, all arms were rewarded and mice simply had to switch arms in order to maximize rewards. Then, after either a 1 or 28 d delay, we either devalued one arm, or we reinforced a specific sequence of pokes. We found that after a 1 d delay mice adapted relatively easily to the changes. In contrast, mice given a 28 d delay struggled to adapt, especially for changes to the optimal sequence of actions. Immediate early gene imaging suggested that the 28 d mice were less reliant on their hippocampus and more reliant on their medial prefrontal cortex. These data suggest that systems consolidation reduces behavioral flexibility, particularly for changes to the optimal sequence of actions.

[Supplemental material is available for this article.]

In a world that is always changing, it is essential for animals to exhibit behavioral flexibility, that is, to adapt their behavior after a change in the environment. Indeed, most species exhibit some flexibility, and can learn to alter their behavioral policies when the world around them changes (Harlow 1949; Rapp 1990; Day et al. 2003; Bond et al. 2007; Asem and Holland 2013). Understanding the neural mechanisms that promote or hinder this behavioral flexibility is crucial to understanding how animals can survive in a dynamic environment (Santoro et al. 2016).

Because we use our memories to guide our actions, changes to our memories may affect our behavioral flexibility. One process that can change memories is consolidation (Müller and Pilzecker 1900). Generally, researchers distinguish between two types of memory consolidation: synaptic consolidation, which helps to stabilize recently formed memories, and systems consolidation, a long-term process by which the circuits involved in the storage and retrieval of memories are altered (Squire and Kandel 2003; Dudai 2004; Frankland and Bontempi 2005; Winocur et al. 2007). A core tenet of many theories of systems consolidation is that some memories that are initially dependent on the hippocampus for storage and retrieval lose this dependency after consolidation. Concomitantly, memories become more dependent on neocortical areas, particularly the medial prefrontal cortex (mPFC) (Zola-Morgan and Squire 1990; Frankland et al. 2004; Vetere et al. 2011; Einarsson and Nader 2012). With respect to the question of behavioral flexibility, systems consolidation may be particularly relevant, as data suggests that it not only alters the substrates on which memories depend, but also transforms their content, rendering them less precise and more “gist-like” (Tse et al. 2007; Wiltgen and Silva 2007; Winocur et al. 2007; Richards et al. 2014; Sweegers et al. 2014). This may have implica-

tions for flexibility, depending upon what is stored by these “gist-like” memories and whether they are themselves changeable or highly static.

However, research regarding the effects of systems consolidation on behavioral flexibility is limited. One previous study examined the impact of systems consolidation on the ability to adapt to new platform locations in the Morris water maze. This study found that systems consolidation can impair flexibility when new platform locations violated previously learned patterns of locations (Richards et al. 2014). Furthermore, some studies have looked at the role of hippocampal plasticity in reversal or latent extinction learning. Latent extinction paradigms require animals to learn passively about devaluation by being confined to an area that was previously rewarded, but with the rewards removed. In contrast, reversal learning paradigms require animals to discover that a different set of actions are now rewarded. These tests of flexibility found that cellular mechanisms for synaptic plasticity in the hippocampus are critical for these forms of flexibility (Gabriele and Packard 2007; Dong et al. 2013). For example, infusion of *N*-methyl-*D*-aspartate (NMDA) receptor agonists into the hippocampus can facilitate latent extinction (Gabriele and Packard 2007), and infusion of pharmacological agents that block long-term depression impair reversal learning (Dong et al. 2013). These studies suggest that the hippocampus may be an important circuit for behavioral flexibility, which carries implications for systems consolidation. But, very little is known about how systems consolidation impacts flexibility in the face of different types of changes to the environment. For example, researchers distinguish between altering the rewards that specific actions provide in isolation and altering the sequence of actions that lead to a reward

**Corresponding author: blake.richards@mcgill.ca**

Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.051243.119>. Freely available online through the *Learning & Memory* Open Access option.

© 2020 Sathiyakumar et al. This article, published in *Learning & Memory*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

(Momennejad et al. 2017). Whether systems consolidation impacts flexibility for these two different forms of change differentially is unknown. This question is important, both because it will help us to understand more about systems consolidation and flexibility, and because it is relevant to questions about the nature of representations used by our memory systems (Stachenfeld et al. 2017).

In order to address this question, we designed a novel Y-maze paradigm that allowed us to explore the impact of systems consolidation on these two types of behavioral flexibility: (1) flexibility when a previously rewarded action is devalued, (2) flexibility when a new optimal sequence of actions is introduced. Mice were trained to obtain rewards in the Y-maze by switching nose pokes between different arms. In the initial training, all arms were rewarded equally and there was no sequential structure to the rewards (as long as a new arm was poked in, there was no advantage to poking in any specific arm). Then after a 1 d (pre-systems-consolidation) or 28 d (post-systems-consolidation) delay we changed the nature of the task and the mice were re-exposed to the Y-maze. As stated above, we introduced two different types of change. In one type of change, we devalued one of the arms, making it optimal to never poke in that arm. In the other type of change, we made it such that certain arms were far more likely to be rewarded after each other, making it optimal to follow a specific sequence of nose pokes. As such, after the delay, the mice had to exhibit behavioral flexibility by changing their actions in order to optimally reap rewards. We found that following a 28 d delay mice were less able to adapt to changes, especially to changes in the optimal sequence of actions. Furthermore, immediate early gene imaging showed reduced activity in the hippocampus and increased activity in mPFC following the 28 d delay. These results suggest that systems consolidation has pronounced impacts on

behavioral flexibility, and in particular, it reduces the ability to flexibly alter the sequence of actions that an animal will take in a given environment. This may relate to changes in the neural substrates underlying the memories.

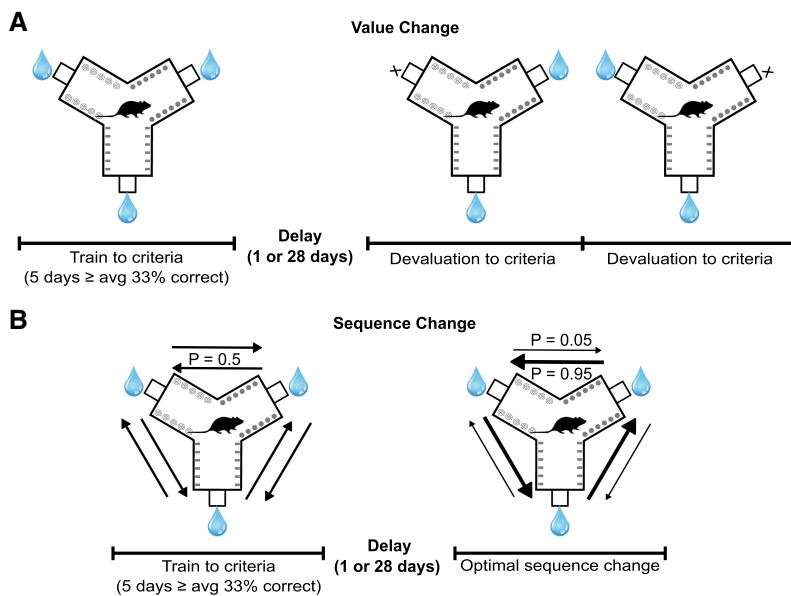
## Results

### A novel Y-maze paradigm for testing different types of behavioral flexibility

We wanted to design a behavioral paradigm that would allow us to distinguish between two important types of behavioral flexibility: (1) flexibility when previously valuable actions are rendered invaluable, (2) flexibility when the optimal sequence of actions is altered. To do this we used a Y-maze where mice could poke at the end of an arm to receive a reward. In order to maximize rewards, the mice could not repeatedly poke in the same arm, and had to switch between arms, similar to other behavioral tasks that have been shown to be hippocampus dependent (Deacon and Rawlins 2006). We could then either devalue one of the arms by inhibiting the delivery of reward from it (Fig. 1A), or we could alter the probability that a reward would move from one arm to another, such that there was a new optimal sequence of pokes introduced (Fig. 1B). The optimal sequence of pokes was changed such that the next reward would be located to the left of the previously rewarded arm with 95% probability and to the right with 5% probability. This design allowed us to independently assess the impact of systems consolidation on flexibility when an action is devalued and flexibility when a new sequence of actions must be learned.

Before changing the task (either by devaluation or sequence change) we trained the mice on a version of the task where all arms were rewarded equally. Once a group of animals reached a criteria of 33% rewarded pokes for 5 d in a row, we would introduce changes to the task in order to probe flexibility. We found that mice were able to learn the initial task well, and after ~1 wk most groups reached our criteria. Mice from different delay groups and flexibility tests did not statistically significantly differ in their performance during this training (Supplemental Fig. S1A repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=5$  mice, (28 d)  $n=7$  mice,  $F_{(1,10)}=2.66$ ,  $P=0.13$ , Supplemental Fig. S1B repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=7$  mice, (28 d)  $n=7$  mice,  $F_{(1,12)}=1.27$ ,  $P=0.28$ ; Supplemental Fig. S6A repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=6$  mice, (28 d)  $n=7$  mice,  $F_{(1,13)}=0.28$ ,  $P=0.60$ , Supplemental Fig. S6B repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=5$  mice, (28 d)  $n=6$  mice,  $F_{(1,11)}=0.04$ ,  $P=0.85$ ).

After training to criteria we then either devalued one arm (Fig. 1A) or altered the sequence probabilities (Fig. 1B), as outlined above. We made the changes to the task following either a 1 or 28 d delay postcriteria being reached. During the delay, no additional training occurred and mice simply rested in their home cages. Previous research has demonstrated that a 28 d delay is sufficient for systems consolidation in mice, whereas a 1 d delay is



**Figure 1.** A novel Y-maze paradigm for testing different types of behavioral flexibility. (A) A Y-maze paradigm to test flexibility when previously valuable actions are rendered invaluable. Mice were first trained on a variation of the task where all arms were rewarded. Once each group of animals successfully achieved 33% rewarded pokes across all trials for five consecutive days the criteria was met and the delay period of 1 or 28 d was initiated. Next, flexibility was probed by inhibiting the release of reward from one arm. Once each group of animals stopped poking in the arm less than three times a day criteria was met and testing ended. (B) A Y-maze paradigm to test flexibility when the optimal sequence of actions is altered. Mice were first trained on a variation of the task where all arms were rewarded. Once each group of animals successfully achieved 33% rewarded pokes across all trials for five consecutive days the criteria was met and the delay period of 1 or 28 d was initiated. Next, flexibility was probed by reinforcing a sequence of pokes for 10 consecutive days.

not (Kim and Fanselow 1992). Thus, the comparison between the two groups would allow us to examine the impact of systems consolidation on flexibility in the face of these two different types of change to the task.

### A prolonged delay partially impairs the ability to adapt to action devaluation

We were interested in understanding how systems consolidation impacts the ability of mice to alter their behavior when a previously rewarded action is no longer valuable. We ran two separate sets of experiments to test the impact of systems consolidation on flexibility to action devaluations. The first set of experiments were an initial test, and the second set of experiments were for the purpose of replication. In addition to the 1 and 28 d delay animals we tested naive animals that had not been trained at all previously, in order to compare the impact of previous learning to situations with no experience.

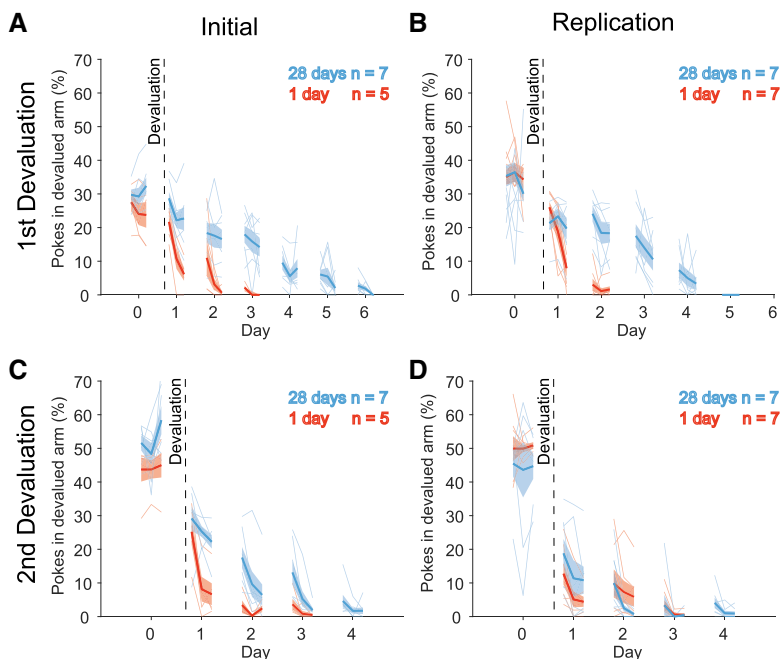
Prior to exposing mice to the tests for flexibility, they were trained on a basic version of the task where all actions were rewarded equally. We found that mice from both delay groups were able to learn the task and they did not perform significantly differently from each other (Supplemental Fig. S1A repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=5$  mice, (28 d)  $n=7$  mice,  $F_{(1,10)}=2.66$ ,  $P=0.13$ , Supplemental Fig. S1B repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=7$  mice, (28 d)  $n=7$  mice,  $F_{(1,12)}=1.27$ ,  $P=0.28$ ).

After the first devaluation we observed that animals that experienced a short 1 d delay rapidly adapted, and stopped poking in the devalued arm after only 2–3 d of being exposed to this variation of the maze. In contrast, the animals in the 28 d delay group

persisted poking in the unrewarded arm for longer, and typically took 4–6 d to stop poking in the unrewarded arm (Fig. 2A; repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=5$  mice, (28 d)  $n=7$  mice,  $F_{(1,10)}=21.13$ ,  $P=9.85 \times 10^{-4}$ ). We found the same result in our replication experiments (Fig. 2B; repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=7$  mice, (28 d)  $n=7$  mice,  $F_{(1,12)}=21.41$ ,  $P=5.8 \times 10^{-4}$ ). We also note that on the last day of training there was no statistical difference in the percentage of pokes in the unrewarded arm between mice from both delay groups (Fig. 2A; repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=5$  mice, (28 d)  $n=7$  mice,  $F_{(1,10)}=2.60$ ,  $P=0.14$ ; Fig. 2B; repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=7$  mice, (28 d)  $n=7$  mice,  $F_{(1,12)}=0.65$ ,  $P=0.43$ ).

Although not statistically significant, the 1 d animals seemed to exhibit a trend of more flexibility than our naive controls, and generally stopped poking in the devalued arm earlier (Supplemental Fig. S2A; repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=5$  mice, (naive)  $n=7$  mice,  $F_{(1,10)}=4.80$ ,  $P=0.05$ ; Supplemental Fig. S2B; repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=7$  mice, (naive)  $n=8$  mice,  $F_{(1,13)}=0.85$ ,  $P=0.38$ ). Moreover, the difference between the 1 and 28 d groups was not simply a result of the 28 d delay animals forgetting the task, because they adapted more slowly than the naive animals with no experience, showing that their previous training was impairing the new learning (Supplemental Fig. S3A; repeated measures ANOVA, percentage of pokes in unrewarded arm: (28 d)  $n=7$  mice, (naive)  $n=7$  mice,  $F_{(1,12)}=4.69$ ,  $P=0.05$ ; Supplemental Fig. S3B; repeated measures ANOVA, percentage of pokes in unrewarded arm: (28 d)  $n=7$  mice, (naive)  $n=8$  mice,  $F_{(1,13)}=22.63$ ,  $P=4.66 \times 10^{-4}$ ). To further rule out forgetting as the cause of inflexibility we examined the percentage of repeated pokes (the prepotent tendency) that the animals made on the first trial of training and the first trial of the postdelay period (testing). We found that both delay groups showed significant decreases in the percentage of repeated pokes between the first ever trial of training and the first trial of the postdelay period (Supplemental Fig. S4; paired *t*-test (1 d, first trial training) versus (1 d, first trial postdelay):  $t_{(11)}=6.47$ ,  $P=4.59 \times 10^{-5}$ ; paired *t*-test (28 d, first trial training) versus (28 d, first trial postdelay):  $t_{(13)}=2.99$ ,  $P=0.01$ ). These results further support the conclusion that the 28 d animals had not completely forgotten their initial training, and fits with the claim that forgetting cannot account for the decreased behavioral flexibility seen in the 28 d delay group.

All of these results were also observed in another set of experiments utilizing a 9 d delay in place of a 28 d delay (Supplemental Fig. S5 repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=7$  mice, (9 d)  $n=7$  mice,  $F_{(1,12)}=68.91$ ,  $P=2.57 \times 10^{-6}$ ). We also performed statistical analyses between the 28 delay groups (from the initial and replication studies) and the 9 d delay groups. This analysis revealed a lack of statistically significant differences in their performance on the action



**Figure 2.** A prolonged delay partially impairs the ability to adapt to action devaluation. (A) In the initial experiment animals that underwent a 28 d delay took longer to cease poking in the first devalued arm. Day zero represents the last day of training before flexibility was probed. (B) In the replication experiment animals that underwent a 28 d delay also took longer to cease poking in the first devalued arm. (C) In the initial experiment animals that underwent a 28 d delay took longer to cease poking in the second devalued arm. (D) In the replication experiment animals from both delay groups took a similar amount of time to cease poking in the second devalued arm. (Bold lines are the mean percentage of pokes for each group, lighter lines are individual animals, and shaded boxes around the mean are the standard error of the mean)

devaluation task (repeated measures ANOVA, percentage of pokes in unrewarded arm: (28 d initial)  $n=7$  mice, (9 d)  $n=7$  mice,  $F_{(1,12)}=0.087$ ,  $P=0.77$ ; repeated measures ANOVA, percentage of pokes in unrewarded arm: (28 d initial)  $n=7$  mice, (9 d)  $n=7$  mice,  $F_{(1,12)}=0.042$ ,  $P=0.84$ ). Together these data suggest that the consolidation process occurs primarily in the first week following learning.

For the second devaluation, we obtained similar results, though there was less consistency between the two experiments. In the initial experiment, the 28 d delay animals continued to show less ability to adapt to the new devaluation than the 1 d delay group (Fig. 2C repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=5$  mice, (28 d)  $n=7$  mice,  $F_{(1,10)}=12.88$ ,  $P=4.99 \times 10^{-3}$ ). However, in the replication experiment we did not see any significant difference between the two delay groups (Fig. 2D; repeated measures ANOVA, percentage of pokes in unrewarded arm: (1 d)  $n=7$  mice, (28 d)  $n=7$  mice,  $F_{(1,12)}=0.16$ ,  $P=0.69$ ). This suggests that the new training during the first devaluation may render the animals more flexible again in some circumstances. Altogether, this data demonstrates that a prolonged delay, potentially permitting systems consolidation, leads to decreased flexibility to adapt to devaluations of a specific action. However, this process does not eliminate the ability to adapt, it only inhibits it.

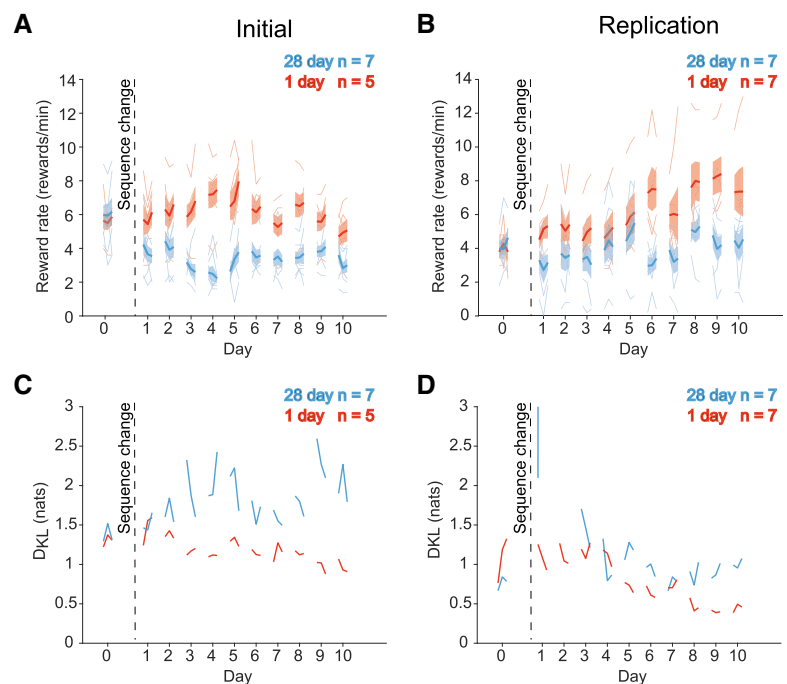
### A prolonged delay completely impairs the ability to adapt to changes in action sequence

We wanted to understand how systems consolidation may impact the ability of mice to alter their behavior when a new optimal action sequence was introduced. As before, we also ran replication experiments to confirm the results. First, mice were trained on the basic version of the task where all arms were rewarded equally. Mice were able to learn this task well and there was no significant difference between the groups' abilities to learn (Supplemental Fig. S6A; repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=6$  mice, (28 d)  $n=7$  mice,  $F_{(1,11)}=0.28$ ,  $P=0.60$ ; Supplemental Fig. S6B; repeated measures ANOVA, percentage of rewarded pokes: (1 d)  $n=5$  mice, (28 d)  $n=6$  mice,  $F_{(1,9)}=0.04$ ,  $P=0.85$ ). As stated above, after the delay, we introduced asymmetric probabilities, such that the arm to the left was rewarded next 95% of the time. Importantly, this new set of transition probabilities provided an explicit optimal strategy (always go to the arm to the left), and as such, we could not only examine the rewards obtained, but also the extent to which the animals matched the optimal strategy in their behavior.

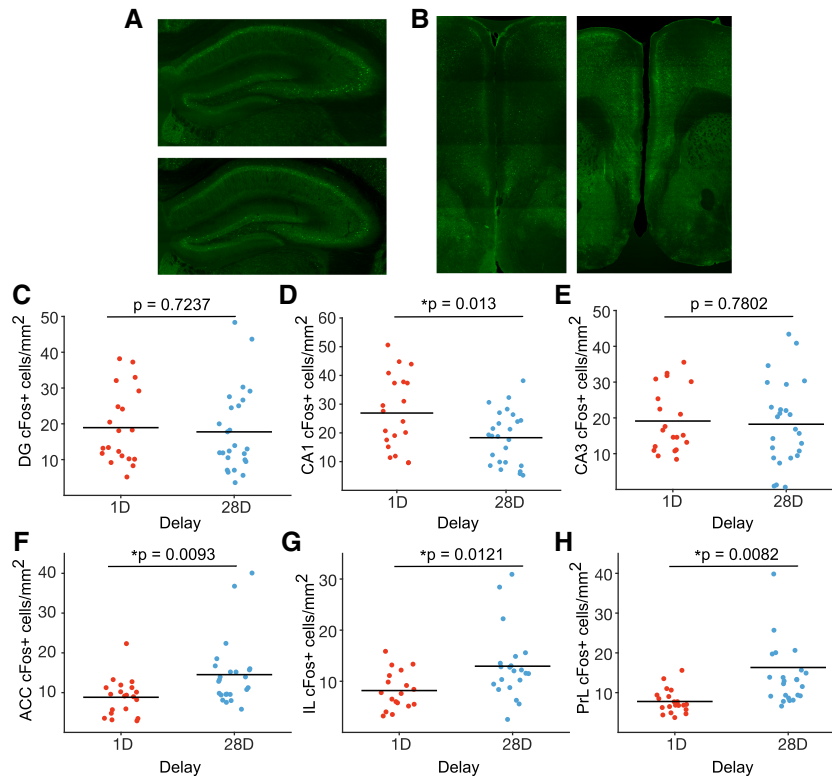
We found that a 28 d delay completely eliminated the ability to adapt to a new optimal sequence. Similar to the action devaluation data, we found that a 28 d delay impaired flexibility when a new optimal sequence was introduced. In this case, the impairment was so pronounced that the 28 d delay animals, never reached the performance levels of the

1 d delay animals within 10 d of testing (Fig. 3A; repeated measures ANOVA, reward rate: (1 d)  $n=6$  mice, (28 d)  $n=7$  mice,  $F_{(1,11)}=14.15$ ,  $P=3.14 \times 10^{-3}$ ). We observed similar results in our replication experiments but they were just barely not statistically significant at the  $\alpha=0.05$  level, most likely due to a cage having to be sacrificed due to poor health (Fig. 3B; repeated measures ANOVA, reward rate: (1 d)  $n=5$  mice, (28 d)  $n=6$  mice,  $F_{(1,9)}=5.23$ ,  $P=0.05$ ). These results were further confirmed by examining the match to the optimal sequence strategy. We measured the Kullback–Leiber divergence ( $D_{KL}$ ) between the groups' behavior and the optimal sequence, and found that the 28 d delay animals were never able to achieve the same degree of match to the optimal strategy that the 1 d delay animals were (Fig. 3C,D).

Again, these results did not appear to be a result of forgetting, as the naive animals were able to learn the optimal sequence as well as the 1 d animals, but only after more time (Supplemental Fig. S7). On the first 3 d after a specific optimal sequence of actions were reinforced naive controls with no prior experience in the maze performed significantly poorly compared to 1 d delay mice (Supplemental Fig. S7A repeated measures ANOVA, reward rate: (1 d)  $n=6$  mice, (naive)  $n=7$  mice,  $F_{(1,11)}=33.8$ ,  $P=1.17 \times 10^{-4}$ ; Supplemental Fig. S7B repeated measures ANOVA, reward rate: (1 d)  $n=5$  mice, (naive)  $n=7$  mice,  $F_{(1,10)}=22.07$ ,  $P=8.43 \times 10^{-4}$ ). Interestingly, the naive controls did not perform significantly different than the 1 d delay animals on the last 3 d of this task (Supplemental Fig. S7A repeated measures ANOVA, reward rate: (1 d)  $n=6$  mice, (naive)  $n=7$  mice,  $F_{(1,11)}=2.30$ ,  $P=0.15$ ;



**Figure 3.** A prolonged delay completely impairs the ability to adapt to changes in action sequence. (A) In the initial experiment 28 d delay animals are unable to reach reward rate levels of the 1 d delay animals. Day zero represents the last day of training before flexibility was probed. (B) In the replication experiment 28 d delay animals are unable to reach reward rate levels of the 1 d delay animals on most days. (C) In these plots the  $D_{KL}$  represents the difference between each delay group's nose-poke strategy and the optimal strategy. The larger the  $D_{KL}$  the further away the group is from the optimal strategy. In the initial experiment animals that underwent a 28 d delay have a higher  $D_{KL}$  than the 1 d delay animals. (D) In the replication experiment animals that underwent a 28 d delay have a higher  $D_{KL}$  on most days meaning that they are further away from the optimal strategy. Bold lines are the mean percentage of pokes for each group, lighter lines are individual animals, boxes around the mean are the standard error, and day zero represents the last day of training before flexibility was probed. Note: The data for the 28 d delay group for day 2 is not presented because they exhibited extremely high  $D_{KL}$ .



**Figure 4.** A prolonged delay alters hippocampal and mPFC activity. (A) Examples of representative cFos stained hippocampal slice (top is from a 1 d delay animal and bottom is from a 28 d delay animal) (B) Examples of cFos stained mPFC slices (left is from a 1 d delay animal and right is from a 28 d delay animal). (C–E) cFos positive cell counts from different regions of the hippocampus. (F–H) cFos positive cell counts from different regions of the mPFC. Individual animals are expressed as solid dots and the horizontal lines are the means for each delay group. Asterisks indicate significant differences following Bonferroni correction.

Supplemental Fig. S7B repeated measures ANOVA, reward rate: (1 d)  $n = 5$  mice, (naive)  $n = 7$  mice,  $F_{(1,10)} = 0.74$ ,  $P = 0.41$ ). In contrast, naive animals performed significantly better than 28 d delay animals during the last 3 d of this task suggesting that consolidation of prior information renders an inability to flexibly adapt to changes to the optimal sequence of actions that should be taken (Supplemental Fig. S8A repeated measures ANOVA, reward rate: (28 d)  $n = 7$  mice, (naive)  $n = 7$  mice,  $F_{(1,12)} = 33.31$ ,  $P = 8.87 \times 10^{-5}$ ; Supplemental Fig. S8B repeated measures ANOVA, reward rate: (28 d)  $n = 6$  mice, (naive)  $n = 7$  mice,  $F_{(1,11)} = 9.28$ ,  $P = 0.01$ ). We also found that the 1 d delay groups showed an almost significant decrease in the percentage of repeated pokes between the first ever trial of training and the first trial of the postdelay period (testing) (Supplemental Fig. S9; paired  $t$ -test (1 d, first trial training) versus (1 d, first trial postdelay):  $t_{(10)} = 2.19$ ,  $P = 0.05$ ). The 28 d delay group showed significant decreases in the percentage of repeated pokes between the first ever trial of training and the first trial of the postdelay period (Supplemental Fig. S9 paired  $t$ -test (28 d, first trial training) versus (28 d, first trial postdelay):  $t_{(12)} = 13.85$ ,  $P = 9.64 \times 10^{-9}$ ). These results further suggest that the 28 d animals had not completely forgotten their initial training. Together, these results suggest that systems consolidation has a particularly strong impact on flexibility in the face of changes to the optimal sequence of actions.

### A prolonged delay alters hippocampal and mPFC activity

We wanted to investigate whether there was any evidence that the underlying memory traces were being altered by the extended de-

lay. Animals in 1 and 28 d delay groups were sacrificed ~90 min after their final trial and hippocampus and mPFC containing slices were stained for cFos as a marker of recent activity (Fig. 4A,B). We counted cFos-positive cells in the dentate gyrus (DG), CA1, and CA3 of the hippocampus and the anterior cingulate (ACC), prelimbic (PrL), and infralimbic cortices (IrL) of the mPFC (Fig. 4C–H). We found that there were reduced cFos positive cells in the CA1 (Fig. 4D  $t$ -test (1 d, CA1) versus (28 d, CA1):  $t_{(43)} = 2.59$ ,  $P = 0.013$ ) region of 28 d delay mice. Furthermore, we found an increased number of cFos-positive cells in the ACC (Fig. 4F  $t$ -test (1 d, ACC) versus (28 d, ACC):  $t_{(42)} = -2.72$ ,  $P = 0.0093$ ), IL (Fig. 4G  $t$ -test (1 d, IL) versus (28 d, IL):  $t_{(37)} = -2.63$ ,  $P = 0.012$ ), and PrL (Fig. 4H  $t$ -test (1 d, PrL) versus (28 d, PrL):  $t_{(42)} = -2.77$ ,  $P = 0.0082$ ) of the mPFC. This change in cFos positive cells between delay groups demonstrates that the memory traces involved in this task are most likely changing over the course of the delay.

### Discussion

We developed a novel y-maze paradigm to test for two specific types of behavioral flexibility: (1) flexibility when previously valuable actions are rendered invaluable and (2) flexibility when the optimal sequence of actions is altered (Fig. 1). We found that a long enough delay for systems consolidation, the process by which the storage and retrieval of memories becomes less dependent on the hippocampus and more dependent on the mPFC (Zola-Morgan and Squire 1990; Frankland et al. 2004; Vetere et al. 2011; Einarsson and Nader 2012), causes impairments to both types of flexibility. Although, animals from both delay groups were able to exhibit flexibility when faced with changes to the value of actions, animals that underwent a 28 d delay exhibited impaired flexibility because they adapted at a slower rate (Fig. 2). Interestingly, we found that animals that underwent a 28 d delay were completely unable to adapt to changes in the optimal sequence of actions whereas, 1 d delay animals were still able to adapt to these changes (Fig. 3). This suggests that a long delay, possibly involving systems consolidation, negatively affects flexibility to action devaluation and completely inhibits flexibility to optimal sequence selection. Furthermore, we found a decrease in cFos positive cells in the CA1 region of the hippocampus and an increase in cFos positive cells in the ACC, IL, and PrL regions of the mPFC of the 28 d delay animals (Fig. 4). It is also noteworthy that statistical analyses performed between a 9 d and 28 d delay groups suggest it's possible that consolidation of action devaluation learning occurs largely in the first week post learning. These data suggest that it is possible that the underlying memory traces are altered, which may be causing the changes in flexibility observed in these experiments.

Theories of complementary learning systems suggests that hippocampal and neocortical structures have different roles in learning and memory. The hippocampus is thought to underlie fast processing of information and the neocortex slowly extracts and stores regularities from this information (McClelland et al.

1995). This is broadly in line with our results as the 28-d delay animals that most likely underwent systems consolidation and possibly relied on memory traces from the mPFC to guide their behavior found it challenging to exhibit flexibility. This could be due to the slow nature of neocortical learning and possibly the decreased contribution of a fast-learning hippocampus in solving these tasks. However, there is evidence that a reminder can render a previously consolidated memory hippocampus-dependent again (Winocur et al. 2009). As such, we must be careful to not assume that the hippocampus was completely disengaged in our 28 d animals. Moreover, we did not find any significant correlations between cFos levels and behavioral flexibility (data not shown), so caution in this interpretation is doubly warranted.

There is another complementary learning system that we did not explore in our study but which deserves some discussion. The striatum and hippocampus have also been thought to work together to solve tasks. The striatum has been implicated in the acquisition of reliable and repetitive sequences of actions or “habit learning” (Knowlton et al. 1996; Voermans et al. 2004; Hartley and Burgess 2005), whereas the hippocampus has been thought to be involved in the expression of flexible and novel responses (Voermans et al. 2004; Hartley and Burgess 2005). Thus, it is important not to rule out the possible contributions the striatum has on the behavioral inflexibility exhibited by the 28 d delay animals. Of course, passive systems consolidation (i.e., without continued training) is not typically considered to involve increased reliance on the striatum, but this cannot be ruled out as a possibility.

Our data suggests that the lack of flexibility exhibited by the 28 d delay animals is not due to forgetting and having to relearn the task. Naive animals with no experience with action devaluation or optimal sequence changes were able to perform better than mice that underwent a long 28 d delay (Supplemental Figs. S3, S8). Also, animals from both delay groups show a decreased percentage of repeated pokes on the first trial of the postdelay period compared to the first trial of the training period (Supplemental Figs. S4, S9). This data touches on an important question: why does the brain engage in systems consolidation if it renders animals less flexible? It has been argued that the transient nature of memories can actually promote flexibility (Richards and Frankland 2017), but of course, not all forgetting is beneficial. The process of systems consolidation likely helps to protect memory traces from forgetting. While this possible protection mechanism could be beneficial in a relatively stable environment it may come at the cost of decreased flexibility.

One potentially interesting implication of our findings is in relation to the question of predictive representations in the hippocampus (Stachenfeld et al. 2017). Theories of a “predictive map” in the hippocampus, utilizing successor representations, predict that hippocampus-dependent learning should be very quick in adapting to devaluations but slow in adapting to changes in sequences (Momennejad et al. 2017). Our data is consistent with this prediction, as the animals adapted more quickly to the devaluation, especially in the 1 d group. However, it should be noted that the change in optimal sequence may simply have been a more difficult task, which could explain why it took the mice longer to adapt to it, even in the 1 d delay group. Future work could further explore these sorts of behavioral distinctions to better test predictive map theories.

The goal of both experimental paradigms used in our study is for animals to maximize rewards in the environment over time. As such, we can think about our study in the context of reinforcement learning. One theory of reinforcement learning suggests that animals learn to associate rewards with latent (unobservable) causes in the environment. This theory predicts that a change in rewards attributed to an already established latent cause will lead an animal’s internal model to update (Gershman et al. 2015), changing

their behavior. In contrast, if an animal attributes a change in reward to a new latent cause they may not alter their behavior. In this framework, it is possible that the 1 d delay animals exhibited flexibility because they updated their internal models, whereas the 28 d delay animals assigned the changes to a new latent cause, leading them to persist in their previously learned behavior.

We speculate that systems consolidation would also have an impact on other training paradigms that require the expression of behavioral flexibility, such as extinction. Our devaluation paradigm is similar in flavor to extinction, as both involve learning that a previously rewarded state will no longer be rewarded. Also, we suspect if the training itself had included changes (e.g., if different arms were devalued throughout training) then the animals would not exhibit this pronounced decline in flexibility as a result of systems consolidation, and may even have been more flexible as their internal “schema” would include the need to change. These are interesting avenues that can be taken into consideration in the future to explore the impact of systems consolidation on flexibility more broadly.

We note that our study was faced with many limitations. First, we took an almost purely behavioral approach to study the possible effect of systems consolidation of flexibility. We intended this study to lay the groundwork for future research that can focus on more specific neural correlates and memory traces involved in (1) flexibility when previously valuable actions are rendered invaluable and (2) flexibility when the optimal sequence of actions is altered. The hippocampus has been widely implicated in spatial maze paradigms, including those that require switching between arms to obtain rewards (Hughes 1965; Crusio et al. 1987; Deacon and Rawlins 2006; Van der Borght et al. 2007) but we did not experimentally confirm here that the hippocampus is required for this task. Furthermore, we use cFos as a correlate for activity, which comes with a host of other caveats. Nonetheless, our cFos data which showed reduced hippocampal activity and increased mPFC activity in the 28 d delay group was very consistent with this being a hippocampus-dependent task that undergoes systems consolidation.

In summary, this work provides initial evidence demonstrating that systems consolidation can impair behavioral flexibility. Notably, systems consolidation seems to impair behavioral flexibility in the face of action devaluations and completely inhibit flexibility to optimal sequence changes. These results provide a framework to understand how our past informs our ability to adapt in the future.

## Materials and Methods

### Animals

All experiments conform to the rules and guidelines of the Canadian Council for Animal Care and the University of Toronto Scarborough Local Animal Care Committee. Wild-type C57BL/6 mice were bred with wild-type 129/SvJ mice (Jackson Laboratories) to produce F1 generation hybrid litters. Adult male hybrid mice (between 8 to 14 wk old) were used in these experiments. Mice were housed in the vivarium on a 12h–12h standard light–dark cycle. They were food restricted to 85% of their free feeding weight and water was available ad-libitum. Training and testing took place during the light cycle between the hours of 10 a.m. and 6 p.m.

### Behavioral apparatus

An automated three armed Y-maze was designed of black Plexiglas with a height of 40 cm, arm length of 20 cm, and arm width 10 cm. Each arm had textural inserts that were custom designed (Solidworks) and printed using a three dimensional (3D) printer (Solidoodle) to provide mice with sensory cues. Each arm was lined

with a different textural insert. At the end of each arm there was a photo interrupter (Sharp GPlA57HRJ00F) used to detect nose pokes. This was slotted into a custom designed 3D printed nose poke port that communicated the poke information to a pinch valve (ASCO 284 Series 17 to 42 mm Solenoid) via an arduino, which controlled delivery of the reward. A custom Matlab script was used to control each pinch valve, the sequence with which rewards were delivered to different ports, and the amount of time the valves remained open. For these experiments, the rewarded pokes resulted in the pinch valve opening for 200 msec and releasing approximately 0.15 mL of Vanilla flavored Boost.

## Handling and habituation

All experiments took place in a dimly lit room with minimal noise. Mice were handled for 5 min by the experimenter and habituated to the behavioral apparatus for five consecutive days prior to training. In order to habituate the mice to the maze a droplet of vanilla flavored Boost was placed in each nose poke port and mice were placed in the middle of the apparatus and allowed to explore for 5 min for five consecutive days.

## Training

Mice were trained to poke in the ports at the end of each tunnel in order to receive a reward via the pinch valve. They were exposed to 5 min trials three times each day. During training, there was an equal chance that each of the three ports could be rewarded. However, the same port was never rewarded twice in a row. In this initial training phase, the probability that any specific port was rewarded after another was even across ports (Fig. 1A,B). As such, the animals simply had to learn to switch between ports after each poke. This was not their natural inclination (which was instead to poke at the same arm many times in a row once a reward had been received). Thus, the initial training taught them to switch between arms. Once each group of animals successfully achieved 33% rewarded pokes across all trials for five consecutive days (Supplemental Fig. S1), the delay period was initiated. The delay period was either 1 or 28 d. After the delay, animals were exposed to altered paradigms (see below).

## Devaluation paradigm

In the devaluation paradigm, one of the three pinch valves was clamped, hence, poking in the port with the clamped valve was no longer a rewarded action. Importantly, though, because there was only a clamp, all other sensory inputs were identical (e.g., the valve still made a tiny sound). As such, the only change from the mouse's perspective was the elimination of any rewards for poking in that arm. Mice were exposed to this variation of the maze 1 or 28 d after training criteria was reached. Mice were again exposed to 5 min trials three times each day.

The mice were trained with a specific arm devalued until they reached a criteria of less than 3 pokes/d in the clamped arm. Once this criteria was met the next day, the clamp was removed from the first arm and a different arm was clamped. Thus, the animal had to learn that rewards would now be delivered in the previously clamped arm and that poking in the newly clamped arm was no longer valuable. Mice were again trained to the same criteria.

## Optimal sequence change paradigm

The optimal sequence of actions needed to reap the most rewards was changed using a custom Matlab script such that the next rewarded port was to the left of the previously rewarded port with 95% probability and to the right with 5% probability. Mice were exposed to this variation of the maze 1 or 28 d after training for 5 min three times a day for 10 consecutive days. To maximize rewards in this paradigm, the mice had to learn to always go to the port to the left.

## Immunohistochemistry

Animals were sacrificed ~90 min after completing the behavioral paradigm. They underwent a transcardial perfusion with PBS and 4% paraformaldehyde (PFA). The brains were carefully extracted, postfixed in 4% PFA for at least 24 h, and sliced to 50  $\mu$ m slices.

Free floating hippocampus and mPFC slices were washed with PBS and incubated in 1% hydrogen peroxide at room temperature. Then slices were blocked in 10% normal goat serum + 0.3% Triton-X-100 in PBS for 2 h at room temperature. Next, slices were incubated overnight at 4°C in a 1:500 dilution of cFos antibody (Santa Cruz, SC-52) in an antibody dilution buffer consisting of 5% normal goat serum + 0.1% Triton-X-100 in PBS.

After primary antibody incubation sections were washed in PBS and incubated in a 1:500 dilution of goat anti-rabbit secondary antibody conjugated with horseradish peroxidase in antibody dilution buffer for 1 h at room temperature. Sections were washed and tyramide signal amplification (TSA) reactions were performed at room temperature in the dark using a TSA plus cyanine 3 system (PerkinElmer, NEL744B001KT). Sections were washed in PBS for 5 min in the dark. Finally, the sections were mounted on slides using fluoroshield (Vector Laboratories) mounting medium to preserve fluorescence and stained for cell nuclei via 4',6-diamidino-2-phenylindole. Two–three slices of the hippocampus and mPFC were imaged per mouse. Images were obtained using a Nikon Eclipse Ni-U epifluorescence scope with a 20 $\times$  objective.

## Data analysis

Behavioral data analysis was automated using custom scripts in Matlab (Mathworks). The script was used to determine the percentage of rewarded pokes, the percentage of pokes in devalued arms, reward rates, and the Kullback–Liebler divergence ( $D_{KL}$ ) between animals' behavior and optimal sequences. The percentage of rewarded pokes was defined as the number of pokes that were rewarded divided by the total number of pokes. The percentage of pokes in the devalued arm was defined as the number of pokes in the devalued arm divided by the total number of pokes in all arms. The reward rate was defined as the total number of rewards received divided by the length of the trial. The  $D_{KL}$  provides a measure of how different two probability distributions are from each other. Hence, we were able to determine the divergence between the distribution of the nose-pokes after flexibility was probed and the optimal distribution of nose-pokes to reap the most rewards. One-way repeated measures analysis of variance (ANOVA) was used to compare averages between 1 and 28 d delay groups and the within subject factors were the animals.

The DG, CA1, and CA3 regions of the hippocampus and the anterior cingulate (ACC), prelimbic (PrL), and infralimbic cortices (Irl) of the mPFC, were manually traced out in Fiji to determine the area in  $\text{mm}^2$ . cFos positive cells of 2–3 slices of the anatomical regions of interest mentioned above were manually quantified using the cell counter plugin on Fiji. To determine the density of cFos positive cells in each region the cell counts were divided by the area (number of cFos positive cells/area  $\text{mm}^2$ ). The densities were averaged for each mouse. Two-tailed, two-sample *t*-tests were used to compare cFos densities between delay groups. Bonferroni correction was used to correct for multiple comparisons.

## Acknowledgments

We would like to thank Paul W. Frankland for helpful comments on an earlier version of this manuscript. This work was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant to B.A.R. (RGPIN-2014-04947).

## References

- Asem JSA, Holland PC. 2013. Immediate response strategy and shift to place strategy in submerged T-maze. *Behav Neurosci* **127**: 854–859. doi:10.1037/a0034686
- Bond AB, Kamil AC, Balda RP. 2007. Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus*, *Nucifraga columbiana*, *Aphelocoma*

- californica*). *J Comp Psychol* **121**: 372–379. doi:10.1037/0735-7036.121.4.372
- Crusio WE, Schwegler H, Lipp H-P. 1987. Radial-maze performance and structural variation of the hippocampus in mice: a correlation with mossy fibre distribution. *Brain Res* **425**: 182–185. doi:10.1016/0006-8993(87)90498-7
- Day LB, Ismail N, Wilczynski W. 2003. Use of position and feature cues in discrimination learning by the whiptail lizard (*Cnemidophorus inornatus*). *J Comp Psychol* **117**: 440–448. doi:10.1037/0735-7036.117.4.440
- Deacon RMJ, Rawlins JNP. 2006. T-maze alternation in the rodent. *Nat Protoc* **1**: 7–12. doi:10.1038/nprot.2006.2
- Dong Z, Bai Y, Wu X, Li H, Gong B, Howland JG, Huang Y, He W, Li T, Wang YT. 2013. Hippocampal long-term depression mediates spatial reversal learning in the Morris water maze. *Neuropharmacology* **64**: 65–73. doi:10.1016/j.neuropharm.2012.06.027
- Dudai Y. 2004. The neurobiology of consolidations, or, how stable is the engraving? *Annu Rev Psychol* **55**: 51–86. doi:10.1146/annurev.psych.55.090902.142050
- Einarsson EO, Nader K. 2012. Involvement of the anterior cingulate cortex in formation, consolidation, and reconsolidation of recent and remote contextual fear memory. *Learn Mem* **19**: 449–452. doi:10.1101/lm.027227.112
- Frankland PW, Bontempi B, Talton LE, Kaczmarek L, Silva AJ. 2004. The involvement of the anterior cingulate cortex in remote contextual fear memory. *Science* **304**: 881–883. doi:10.1126/science.1094804
- Frankland PW, Bontempi B. 2005. The organization of recent and remote memories. *Nat Rev Neurosci* **6**: 119–130. doi:10.1038/nrn1607
- Gabriele A, Packard MG. 2007. D-Cycloserine enhances memory consolidation of hippocampus-dependent latent extinction. *Learn Mem* **14**: 468–471. doi:10.1101/lm.528007
- Gershman SJ, Norman KA, Niv Y. 2015. Discovering latent causes in reinforcement learning. *Curr Opin Behav Sci* **5**: 43–50. doi:10.1016/j.cobeha.2015.07.007
- Harlow HF. 1949. The formation of learning sets. *Psychol Rev* **56**: 51–65. doi:10.1037/h0062474
- Hartley T, Burgess N. 2005. Complementary memory systems: competition, cooperation and compensation. *Trends Neurosci* **28**: 169–170. doi:10.1016/j.tins.2005.02.004
- Hughes KR. 1965. Dorsal and ventral hippocampus lesions and maze learning: influence of preoperative environment. *Can J Psychol* **19**: 325–332. doi:10.1037/h0082915
- Kim J, Fanselow M. 1992. Modality-specific retrograde amnesia of fear. *Science* **256**: 675–677. doi:10.1126/science.1585183
- Knowlton BJ, Mangels JA, Squire LR. 1996. A neostriatal habit learning system in humans. *Science* **273**: 1399–1402. doi:10.1126/science.273.5280.1399
- McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* **102**: 419–457. doi:10.1037/0033-295X.102.3.419
- Momennejad I, Russek EM, Cheong JH, Botvinick MM, Daw ND, Gershman SJ. 2017. The successor representation in human reinforcement learning. *Nat Hum Behav* **1**: 680–692. doi:10.1038/s41562-017-0180-8
- Müller GE, Pilzecker A. 1900. Experimentelle beiträge zur lehre vom gedächtnis. *Z Psychol Ergänzungsband* **1**: 1–300.
- Rapp PR. 1990. Visual discrimination and reversal learning in the aged monkey (*Macaca mulatta*). *Behav Neurosci* **104**: 876–884. doi:10.1037/0735-7044.104.6.876
- Richards BA, Frankland PW. 2017. The persistence and transience of memory. *Neuron* **94**: 1071–1084. doi:10.1016/j.neuron.2017.04.037
- Richards BA, Xia F, Santoro A, Husse J, Woodin MA, Josselyn SA, Frankland PW. 2014. Patterns across multiple memories are identified over time. *Nat Neurosci* **17**: 981–986. doi:10.1038/nn.3736
- Santoro A, Frankland PW, Richards BA. 2016. Memory transformation enhances reinforcement learning in dynamic environments. *J Neurosci* **36**: 12228. doi:10.1523/JNEUROSCI.0763-16.2016
- Squire LR, Kandel ER. 2003. *Memory: from mind to molecules* (Vol. 69). Macmillan, New York.
- Stachenfeld KL, Botvinick MM, Gershman SJ. 2017. The hippocampus as a predictive map. *Nat Neurosci* **20**: 1643–1653. doi:10.1038/nn.4650
- Sweegers CCG, Takashima A, Fernández G, Talamini LM. 2014. Neural mechanisms supporting the extraction of general knowledge across episodic memories. *Neuroimage* **87**: 138–146. doi:10.1016/j.neuroimage.2013.10.063
- Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RGM. 2007. Schemas and memory consolidation. *Science* **316**: 76–82. doi:10.1126/science.1135935
- Van der Borght K, Havekes R, Bos T, Eggen BJL, Van der Zee EA. 2007. Exercise improves memory acquisition and retrieval in the Y-maze task: relationship with hippocampal neurogenesis. *Behav Neurosci* **121**: 324–334. doi:10.1037/0735-7044.121.2.324
- Vetere G, Restivo L, Cole CJ, Ross PJ, Ammassari-Teule M, Josselyn SA, Frankland PW. 2011. Spine growth in the anterior cingulate cortex is necessary for the consolidation of contextual fear memory. *Proc Natl Acad Sci* **108**: 8456–8460. doi:10.1073/pnas.1016275108
- Voermans NC, Petersson KM, Daudey L, Weber B, van Spaendonck KP, Kremer HPH, Fernández G. 2004. Interaction between the human hippocampus and the caudate nucleus during route recognition. *Neuron* **43**: 427–435. doi:10.1016/j.neuron.2004.07.009
- Wiltgen BJ, Silva AJ. 2007. Memory for context becomes less specific with time. *Learn Mem* **14**: 313–317. doi:10.1101/lm.430907
- Winocur G, Moscovitch M, Sekeres M. 2007. Memory consolidation or transformation: context manipulation and hippocampal representations of memory. *Nat Neurosci* **10**: 555–557. doi:10.1038/nrn1880
- Winocur G, Frankland PW, Sekeres M, Fogel S, Moscovitch M. 2009. Changes in context-specificity during memory reconsolidation: selective effects of hippocampal lesions. *Learn Mem* **16**: 722–729. doi:10.1101/lm.1447209
- Zola-Morgan S, Squire L. 1990. The primate hippocampal formation: evidence for a time-limited role in memory storage. *Science* **250**: 288–290. doi:10.1126/science.2218534

Received November 29, 2019; accepted in revised form February 11, 2020.