Retrieval as a Fast Route to Memory Consolidation

Article *in* Trends in Cognitive Sciences · June 2017

DOI: 10.1016/j.tics.2017.05.001

James Antony
Princeton University

Catarina S Ferreira
University of Birmingham

Maria Wimber
University of Birmingham

valuation and affect (mPFC). Their findings extend earlier work illustrating the role of the mPFC in mediating popularity [2], revealing that the mPFC likely has a more global role in indexing social evaluation processes.

Given that participants were scanned while only passively viewing other group members – rather than performing a task that required them to proactively think about social bonds within their network (such as choosing whether to cooperate with a given member) – it is particularly striking that the authors were able to identify neural systems that tracked information about each group member's social standing (i.e., their social status, distance, and ability to broker information). That we appear to do these computations spontaneously, rather than on a need-to-know basis, suggests that our brains are always preparing for social engagement with each individual, regardless of whether such engagement is needed.

These findings provide insight into a question that has been the topic of much theoretical debate in recent years: how do we encode another individual's social value? [3]. The authors' finding that social status is indexed by the mPFC, a region that has a critical role in computing value across numerous decision-making contexts [3], indicates that the mPFC represents the social value of another individual even before deciding to trust, help, or cooperate with that person. However, it is unclear why this region would spontaneously represent these social values in situations where no choice is required. One possibility is that the mPFC response is laying the groundwork for a subsequent decision, signaling a potential opportunity for bolstering one's own social status by connecting with a high-status individual.

Indeed, in conjunction with classic evolution theory [4,5], Parkinson and colleagues' findings suggest that the brain is pre-emptively evaluating other individuals to strategically bias subsequent encounters [12]. One way to interrogate how this value signal is being operationalized would be to measure if there is an association between the integrity of the mPFC response and subsequent decisions to abuse trust, attenuate cooperative actions, or only offer help when there is something to be gained socially. If our brains are keeping track of group members' social status and distance, not only should prosocial acts be exhibited more readily towards a person of high status within the network, which should scale with the mPFC response (compared with those who exhibit lower interconnectedness and status), but these decisions should also be made more quickly and reflexively. Additionally, given the impact of empathy and theory of mind capacities on decisions to trust and be altruistic [6,7], prosocial behavior may be further biased by the interaction between an individual's empathic ability and each group members' social status, distance, and brokerage.

Intriguingly, these findings may also provide insight into the motivations behind punitive behavior, something that remains somewhat elusive [8–11]. One possibility is that doling out punishment is contingent on where the perpetrator stands within the social network. For example, the degree of punishment may parametrically scale with the perpetrator's social status and distance from the individual conferring the punishment. This would suggest that the utility of punishment depends, in part, on how well situated a perpetrator is within the social community.

Parkinson and colleagues [12] have successfully measured how the brain encodes real-world social connections, demonstrating that we track in real time not only where we stand among our peers, but also how our peers measure up against everyone else. This innovative work adds to a budding literature characterizing how we understand our social worlds, while also bringing to the forefront further questions about how we plan to interact with others who occupy our world.

[1]Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI 02906, USA

*Correspondence:
oriel_feldmanhall@brown.edu (O. FeldmanHall).

http://dx.doi.org/10.1016/j.tics.2017.05.005

References

1. Dunbar, R.I.M. (2014) The social brain psychological underpinnings and implications for the structure of organizations. *Curr. Dir. Psychol. Sci.* 23, 109–114
2. Zerubavel, N. *et al.* (2015) Neural mechanisms tracking popularity in real-world social networks. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15072–15077
3. Ruff, C.C. and Fehr, E. (2014) The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* 15, 549–562
4. de Waal, F. (1982) *Chimpanzee Politics: Power and Sex among Apes,* Harper and Row
5. Clutton-Brock, T. (2009) Cooperation between non-kin in animal societies. *Nature* 462, 51–57
6. Masten, C.L. *et al.* (2011) An fMRI investigation of empathy for 'social pain' and subsequent prosocial behavior. *Neuroimage* 55, 381–388
7. FeldmanHall, O. *et al.* (2015) Empathic concern drives costly altruism. *Neuroimage* 105, 347–356
8. Carlsmith, K.M. *et al.* (2002) Why do we punish? Deterrence and just deserts as motives for punishment. *J. Pers. Soc. Psychol.* 83, 284–299
9. Fehr, E. and Schmidt, K.M. (1999) A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114, 817–868
10. Dreber, A. *et al.* (2008) Winners don't punish. *Nature* 452, 348–351
11. FeldmanHall, O. *et al.* (2014) Fairness violations elicit greater punishment on behalf of another than for oneself. *Nat. Commun.* 5, 5306
12. Parkinson, C. *et al.* (2017) Spontaneous neural encoding of social network position. *Nat. Hum. Behav.* 1, 0072

Forum

Retrieval as a Fast Route to Memory Consolidation

James W. Antony,[1,*] Catarina S. Ferreira,[2] Kenneth A. Norman,[1] and Maria Wimber[2,*]

Retrieval-mediated learning is a powerful way to make memories last, but its neurocognitive mechanisms remain unclear. We propose that retrieval acts as a rapid consolidation event, supporting the creation of adaptive hippocampal–neocortical representations via the 'online' reactivation of

associative information. We describe parallels between online retrieval and offline consolidation and offer testable predictions for future research.

Introduction

For over a century, psychologists have known that repeatedly and actively retrieving information from memory, as opposed to restudying the same information, strongly enhances long-term retention [1]. The benefits of retrieval-mediated learning (also known as the 'testing effect') hold across a wide variety of materials and testing formats and remain evident across much of the lifespan [1]. No mechanistic framework exists to date integrating these behavioral findings with the growing literature on the neural basis of learning and memory [1]. Here we attempt such an explanation.

In short, we propose that retrieval acts as a fast route to memory consolidation. Specifically, we propose that retrieval integrates the memory with stored neocortical knowledge and differentiates it from competing memories, thereby making the memory less hippocampus dependent and more readily accessible in the future. We explore theoretical links between retrieval and offline consolidation, describe some key evidence in support of a shared mechanism, draw parallels between this proposal and other forms of rapid consolidation, and outline predictions for future research (Box 1).

Retrieval as a Rapid Consolidation Event

Extensive evidence from rodents and amnesia patients shows that hippocampal damage affects the formation of new declarative memories while leaving remote memories (at least partially) intact [2]. An influential computational model [Complementary Learning Systems (CLS)] [2] suggests that the hippocampus and neocortex act synergistically to allow new learning while preserving old information. Specifically, the neocortex learns slowly and specializes in storing the statistical structure of experiences. The hippocampus learns quickly and specializes in rapidly encoding and binding together new cortical associations. Repeated interactions between the two systems allow new information to slowly shape neocortical representations. If the hippocampus is damaged before enough hippocampal–neocortical interactions can occur, long-term memory will be impaired. These ideas constitute systems-level consolidation, or the process by which newly acquired information is transformed into a stable, long-term memory representation [3].

The gradual transformation that a memory undergoes during systems-level consolidation is promoted by the memory's repeated offline reactivation ('replay') in hippocampal–neocortical circuits. Reactivations during non-rapid eye movement (NREM) sleep arguably play a unique role in embedding information in the neocortex, facilitated by low cholinergic activity and coordinated oscillatory interactions between the hippocampus and neocortex [4]. Replay occurring during both post-learning wakeful rest and sleep has been shown to enhance memory retention [3,4]. Critically, we propose that the neural reactivation of recently acquired memories, as triggered online by incomplete reminders (pattern completion), promotes long-term retention in a way similar to offline replay.

We argue that retrieval and sleep can qualitatively transform memories in at least two distinct ways: by integrating new memories into preexisting neocortical knowledge structures and by adaptively differentiating memories (i.e., reducing their neural overlap) so as to minimize competition between overlapping memories. From a computational perspective, both the integration and differentiation effects can be explained by the tendency of retrieval to be imprecise; that is, to coactivate memories that are semantically or episodically linked to the target memory [5]. Repeated imprecise reactivations in hippocampal–neocortical circuits afford an opportunity to integrate an initially hippocampus-dependent memory into the coactivated neocortical knowledge structures, similar to replay events during NREM sleep (Figure 1). According to [5], the nature of learning driven by coactivation depends on how strongly memories are activated: strong coactivation of memories leads to integration of those memories, whereas moderate activation of competing memories triggers their adaptive weakening [6] and pushes retrieved and competing memories apart in representational space [7], leaving the retrieved memory in a distinct, accessible state for future recall (Figure 1). Importantly, restudy (i.e., simple re-exposure to a complete, previously stored memory) does not share these computational characteristics [6,7]. Restudy may re-impose the memory's original pattern onto the hippocampus and neocortex, causing some strengthening of the original trace. However, because restudy triggers less coactivation of related memories it does not adaptively shape the hippocampal–neocortical memory landscape in the same way as active retrieval.

So far we have outlined the similarities between retrieval and offline consolidation on a theoretical, computational level. We next examine behavioral and neural parallels between memory retrieval effects (indexed by differences in retrieved versus restudied information) and consolidation effects (indexed by sleep versus wake intervals) that empirically support our rapid consolidation view.

Similarities between Retrieval and Consolidation

If retrieval rapidly embeds a memory in the neocortex, future retrievals of this memory can utilize neocortical in addition to hippocampal representations to access the memory. Retrieval-mediated memory boosts should thus be most evident whenever hippocampal traces are weak and recall is relatively more dependent on the neocortex. Consistent with this notion, testing effects are strongest at long delays of several days to weeks

[1], when recall supposedly relies more heavily on neocortical traces than at short delays.

Online retrieval and offline reactivation also share the property of protecting memories from retroactive interference. Including a period of sleep (as opposed to wake) between two learning sessions reduces retroactive interference [3], presumably due to offline reactivation. Intriguingly, including retrieval practice between the learning of two lists of words similarly reduces retroactive interference [8], consistent with a rapid online consolidation mechanism.

Sleep has been shown to globally enhance long-term retention but seems to preferentially benefit salient, rewarded, or emotional information containing future utility [3,4]. Assuming that prioritized information is 'tagged' to receive greater offline processing during sleep, retrieval could in theory interact with the consolidation process in two different ways. If retrieval effectively tags information to receive greater subsequent offline processing during sleep, retrieved memories should benefit more from sleep than restudied memories. By contrast, if the benefits of retrieval occur online, as proposed by our rapid consolidation framework, sleep should offer greater relative benefits to restudied information because retrieved information has already been consolidated online. In fact, restudied information does benefit more from sleep than retrieved information [9], underscoring the idea that sleep offers relatively more benefits to information that has not already been stabilized via online consolidation during retrieval.

Systems consolidation accounts [2] state that initially hippocampus-dependent memories become gradually incorporated into the neocortex for stable long-term storage. If retrieval accelerates consolidation, it should produce neural changes

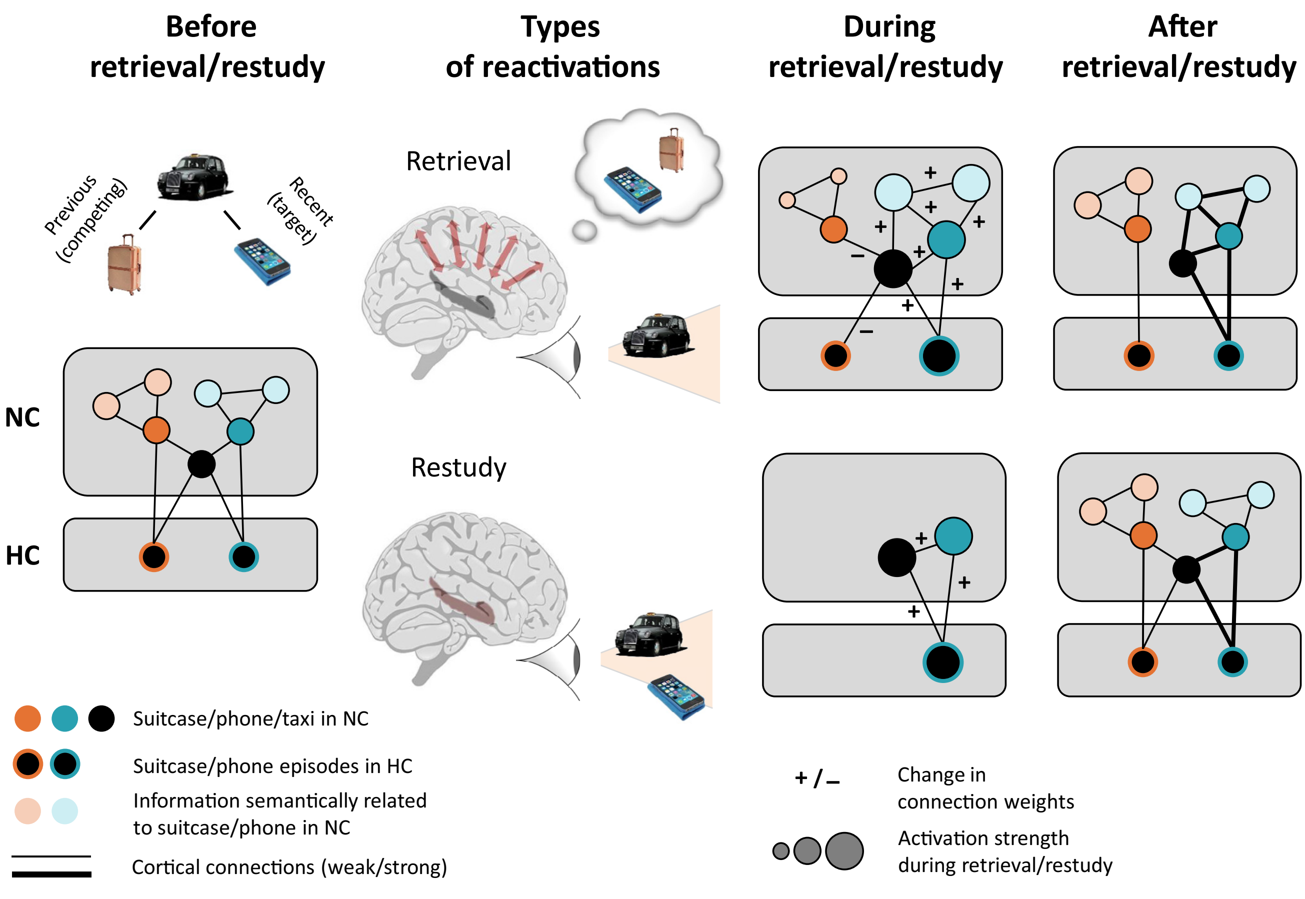

Figure 1. Neural Network Changes via Retrieval Compared with Restudy. Imagine you form a new memory of leaving your phone in a taxicab. Also imagine that 2 weeks before this episode you needed a taxi because you had a very heavy suitcase with you. The two episodes have distinct, nonoverlapping representations in the hippocampus (HC) but are connected via the shared element 'taxi' in the neocortex (NC). The NC also holds representations of memories that are loosely connected with phones and suitcases. Retrieval of the taxi–phone episode will activate the core representations in the HC and NC but will also cause coactivation of related information including the suitcase and knowledge related to phones. Strongly active memories and their connections (blue elements) will be strengthened (indicated by +), supporting memory integration. Moderately coactivated information will be weakened (indicated by –), supporting differentiation, while no change in connection strength occurs for weakly active memories. By comparison, restudy reinstantiates the cortical representations of only those elements that are explicitly presented; connections between these elements will be strengthened but the memory will not be integrated with or differentiated from other stored knowledge.

similar to those with offline consolidation. In humans, connectivity between the hippocampus and the ventromedial prefrontal cortex (vmPFC), a key region in systems consolidation, increases after sleep [10] and retrieval practice but not after restudy [11]. Similarly, it has been shown in rodents that repeated reactivations support the creation of memory traces that can be accessed independently of the hippocampus [12], suggesting that online reactivation resembles offline consolidation at a neural level.

Can Neocortical Representations Be Modified Rapidly?

While most frameworks propose that systems consolidation happens slowly via hippocampal–neocortical interactions over time [2], we propose that retrieval promotes the rapid development of neocortical representations without time and sleep. Importantly, two other learning paradigms suggest that rapid cortical learning is possible. First, research on schema-based learning demonstrates that information consistent with prior knowledge (schemas) can become hippocampus independent and embedded in the neocortex surprisingly quickly and without intervening sleep [13]. As argued above, we assume that retrieval rapidly consolidates memories in a similar way, by coactivating related knowledge structures to facilitate neocortical integration.

Second, in a related paradigm called 'fast mapping' individuals rely on prior knowledge to learn new information by active inference [14]. Intriguingly, patients with hippocampal amnesia can learn new objects more effectively using fast mapping than a simple instruction to encode explicitly [14], suggesting that fast mapping arises from a faster neocortical embedding process [15]. Associations learned via fast mapping (compared with explicit encoding) also benefit less from sleep [15], very similar to associations learned via active retrieval (compared with restudy) [9]. These findings imply that offline enhancement is reduced if cortical integration occurs online via fast mapping or retrieval. Indeed, fast mapping and retrieval share common elements of active inference and knowledge activation and may thus exert their beneficial effects on long-term retention via the same mechanism.

Box 1. Predictions and Open Questions for Future Research

- Human fMRI studies should find a relative increase of vmPFC and decrease of hippocampal contributions when probing previously retrieved compared with restudied memories.
- Retrieval benefits (i.e., testing effects) should be associated with representational differentiation and competitor weakening [5,6].
- Offline reactivation can occur spontaneously but can also be induced by presenting cues previously paired with learning [in a technique known as targeted memory reactivation (TMR)] [4]. As cognitive control is absent during sleep, does TMR result in coactivation of related associations in a manner resembling retrieval during wake [5,6]?
- If repeated reactivations via retrieval help to embed memories in the neocortex and make them less hippocampus dependent, this transformation should have observable effects on behavior. For example, retrieval, like sleep, should promote insight, inference, and generalization, more so than restudy [4], and may reduce context dependency during future recall.
- Interactions between the hippocampus and neocortex should be most effective when capitalizing on preexisting neocortical schemas [13]. If retrieval-mediated learning depends on such interactions, its benefits will be reduced for novel materials that do not have preexisting neocortical representations (e.g., nonsense drawings or syllables).

Concluding Remarks

We propose that retrieval stabilizes memories via mechanisms similar to those that occur during sleep and offline consolidation periods. We acknowledge that sleep has unique characteristics that are not shared with wake retrieval modes. At present it is unclear to what degree reactivations during sleep are imprecise and associated with the same strengthening and weakening dynamics as reactivations during wake. Nevertheless, the rapid consolidation view of retrieval makes testable predictions for future research (Box 1) and will hopefully stimulate new research to bridge the existing gap between cognitive and neuroscientific investigations of memory modification.

Acknowledgments

This work was supported by the CV Starr Fellowship (awarded to J.W.A.), NIMHR01MH069456 (awarded to K.A.N.), and ESRC grant ES/M001644/1 (awarded to M.W.).

[1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA
[2]School of Psychology, University of Birmingham, Birmingham, B15 2TT, UK

*Correspondence:
jantony@princeton.edu (J.W. Antony) and
m.wimber@bham.ac.uk (M. Wimber).

http://dx.doi.org/10.1016/j.tics.2017.05.001

References

1. Rowland, C.A. (2014) The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol. Bull.* 140, 1432–1463
2. McClelland, J.L. *et al.* (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
3. Dudai, Y. *et al.* (2015) The consolidation and transformation of memory. *Neuron* 88, 20–32
4. Antony, J.W. and Paller, K.A. (2017) Hippocampal contributions to declarative memory consolidation during sleep. In *The Hippocampus from Cells to Systems* (Hannula, D.E. and Duff, M.C., eds), pp. 245–280, Springer
5. Norman, K.A. *et al.* (2007) A neural network model of retrieval-induced forgetting. *Psychol. Rev.* 114, 887–953
6. Wimber, M. *et al.* (2015) Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nat. Neurosci.* 18, 582–589
7. Hulbert, J.C. and Norman, K.A. (2015) Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cereb. Cortex* 25, 3994–4008
8. Potts, R. and Shanks, D.R. (2012) Can testing immunize memories against interference? *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 1780–1785
9. Bäuml, K.-H.T. *et al.* (2014) Sleep can reduce the testing effect: it enhances recall of restudied items but can leave recall of retrieved items unaffected. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 1568–1581
10. Gais, S. *et al.* (2007) Sleep transforms the cerebral trace of declarative memories. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18778–18783
11. Wing, E.A. *et al.* (2013) Neural correlates of retrieval-based memory enhancement: an fMRI study of the testing effect. *Neuropsychologia* 51, 2360–23670
12. Lehmann, H. *et al.* (2009) Making context memories independent of the hippocampus. *Learn. Mem.* 16, 417–420
13. Tse, D. *et al.* (2007) Schemas and memory consolidation. *Science* 316, 76–82
14. Sharon, T. *et al.* (2011) Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 108, 1146–1151
15. Himmer, L. *et al.* (2016) Sleep-mediated memory consolidation depends on the level of integration at encoding. *Neurobiol. Learn. Mem.* 137, 101–106