4750 • The Journal of Neuroscience, June 2, 2021 • 41(22):4750–4753

Journal Club

Editor's Note: These short reviews of recent *JNeurosci* articles, written exclusively by students or postdoctoral fellows, summarize the important findings of the paper and provide additional insight and commentary. If the authors of the highlighted article have written a response to the Journal Club, the response can be found by viewing the Journal Club at www.jneurosci.org. For more information on the format, review process, and purpose of Journal Club articles, please see http://jneurosci.org/content/jneurosci-journal-club.

Challenges in Understanding the Role of Reactivation in Modifying Hippocampal Representations

Hannah Tarder-Stoll, Camille Gasser, Wangjing Yu, and Halle R. Dimsdale-Zucker
Department of Psychology, Columbia University, New York, New York 10027
Review of Molitor et al.

Introduction

Our daily lives consist of many repeated events. This poses a challenge for the memory system: because our experiences may contain related elements, we often have to distinguish between similar memories (O'Reilly and McClelland, 1994). The brain can help mitigate interference between memories by driving neural representations of related experiences further apart, a process called differentiation (Schlichting et al., 2015; Favila et al., 2016; Chanales et al., 2017), or by pulling them closer together, a process called integration (Schlichting and Preston, 2015). Empirical and computational work has demonstrated that different hippocampal subfields tend to play different roles in these processes: the dentate gyrus, CA_2, and CA_3 subfields ($DG/CA_{2,3}$) often differentiate related memories, whereas CA_1 often integrates them (Schapiro et al., 2017; Dimsdale-Zucker et al., 2018; Duncan and Schlichting, 2018). One factor that may influence whether the neural representations of overlapping events are integrated or differentiated is the degree to which previous, related experiences are reactivated during new learning (Schlichting et al., 2015; Kim et al., 2017; Ritvo et al., 2019; Wammes et al., 2021). How varying levels of reactivation impact memory representations in different hippocampal subfields, in light of their known representational biases, remains an open question.

To address this, Molitor et al. (2021) used fMRI to characterize the relationship between reactivation and representations of overlapping experiences across hippocampal subfields $DG/CA_{2,3}$, CA_1, and subiculum. Participants first learned AB associations [a particular face or house (A) paired with a particular shape (B)]. They then encoded partially overlapping BC associations [a previously learned shape (B) paired with a new object (C)]. To explore how variation in reactivation strength alters memory representations, the authors assessed the relationship between (1) the strength of reactivation of A-item representations during overlapping BC learning, and (2) how much the similarity between A and C item representations that were linked by a shared B shape changed from before to after encoding. Here, similarity refers to the correlation between multivariate brain activity patterns elicited while participants viewed A and C items both before and after learning (Kriegeskorte, 2008). Increased similarity between these activity patterns after learning suggests that A and C representations overlapped more after being linked to a shared B associate, and was thus taken as evidence of integration. Decreased similarity was taken as evidence of differentiation (Schlichting et al., 2015).

Of note, the authors also varied the perceptual similarity of the shared B items between overlapping item pairs, such that some AB/BC triads contained identical B shapes, while others contained B shapes with varying levels of perceptual similarity (high, low, or no similarity). Changes in memory representations within hippocampal subfields were measured only for AB/BC triads in which the similarity between B shapes was high. Finally, participants were tested on their ability to infer the indirect association between A and C items, allowing the authors to explore the behavioral consequences of reactivation and subsequent changes in item representations.

Ultimately, Molitor et al. (2021) found that when memories for A items were strongly reactivated in the cingulate, occipital, and parietal cortex during BC learning, $DG/CA_{2,3}$ and subiculum differentiated memory representations (i.e., lower similarity between A- and C-item brain activity patterns after, relative to before, learning), whereas CA_1 integrated them (i.e., greater similarity between A- and C-item representations after learning). However, when memories were weakly reactivated in the cortex, $DG/CA_{2,3}$ integrated memory representations. The strength of memory reactivation in the cortex and of integration in the subiculum was also positively associated with behavioral performance on the indirect association test. Together, these findings

Received Feb. 12, 2021; revised Mar. 25, 2021; accepted Apr. 25, 2021.

We thank Mariam Aly, Christopher Baldassano, Lila Davachi, Serra Favila, Zachariah Reagh, and Margaret L. Schlichting as well as the authors of the original Molitor et al. (2021) paper for helpful comments on the manuscript.

Correspondence should be addressed to Hannah Tarder-Stoll at hrt2116@columbia.edu.

https://doi.org/10.1523/JNEUROSCI.0334-21.2021

reveal that cortical reactivation can lead to divergent representation changes in hippocampal subfields and have behavioral consequences for flexible memory use. An open question, however, is how the observed relationship between reactivation and the modification of memory representations arises.

Intrinsic Properties of Hippocampal Subfields Shape Reactivation and Bias Representations

One way to understand how different hippocampal subfields show distinct changes in memory representations is by considering their unique firing properties. Specifically, the innate excitability of each subfield might constrain the magnitude of memory reactivation elicited during BC learning. The strength of this reactivation, in turn, may dictate how overlapping memory representations are modified. Following the predictions of the nonmonotonic plasticity hypothesis (NMPH), weak levels of reactivation will leave memories unchanged, moderate levels of reactivation will lead to differentiation, and strong levels will lead to integration (Ritvo et al., 2019). Let us consider how this might play out across different subfields: CA_1 has relatively dense activity and low inhibition, resulting in a greater likelihood of correlated firing across populations of neurons (Kesner and Rolls, 2015). As such, activation of the newly formed BC representation in CA_1 during learning may be more likely to trigger reactivation of the related, pre-existing AB representation. This strong coactivation of the new BC and existing AB pairs would then promote integration of the two memories, given that correlated firing should strengthen the connections between the neural ensembles representing each individual memory (Ritvo et al., 2019).

In contrast, DG is characterized by high inhibition and sparse neural coding (Kesner and Rolls, 2015). Activation of the BC representation would thus be less likely to initiate strong reactivation of the preexisting AB memory trace. Consequently, these representations are unlikely to become integrated, and may even become differentiated—a claim supported by work suggesting that relatively moderate coactivation of related memory representations weakens the neural connections between them (Ritvo et al., 2019; Wammes et al., 2021). Therefore, the intrinsic properties of DG and CA_1 may drive the dissociable changes in hippocampal representations that Molitor et al. (2021) observed.

The properties of DG alone, however, cannot explain why DG/$CA_{2,3}$ differentiated when cortical reactivation was strong but integrated when reactivation was weak (Molitor et al., 2021). One way to reconcile this finding with the general bias of DG toward differentiation is to consider the complications introduced by combining DG with $CA_{2,3}$, as is typical in human fMRI work. In contrast to the sparse coding scheme and high inhibition of DG, CA_3 is characterized by dense excitatory recurrent connections, which promote memory reactivation (Kesner and Rolls, 2015). Although some studies show differentiation in DG/$CA_{2,3}$ (Dimsdale-Zucker et al., 2018), others find that CA_3 can integrate memories (Grande et al., 2019), indicating that representations in this region may be flexible (Yassa and Stark, 2011; Aly and Turk-Browne, 2018). Ultimately, this flexibility in how memory representations are modified within CA_3, in addition to its distinct firing properties relative to DG, may jointly explain why the combined DG/$CA_{2,3}$ region in Molitor et al. (2021) exhibits integration in some cases, and differentiation in others.

In contrast to these well characterized dynamics of CA_1, CA_3, and DG, less is known about the role of subiculum in differentiation and integration. Anatomically, the subiculum receives substantial input from CA_1, but not CA_3 (O'Mara et al., 2001). As such, one might predict that representations in subiculum would mirror those in CA_1 (Bakker et al., 2008). However, Molitor et al. (2021) find evidence of differentiation in subiculum at strong levels of reactivation, similar to DG/$CA_{2,3}$. Given the relative dearth of knowledge about the role of subiculum in modifying overlapping memory traces, this finding could motivate new exploration into the representation biases of this region.

The Role of Bidirectional Hippocampal–Cortical Communication in Memory Reactivation

Thus far, we have discussed how intrinsic properties of hippocampal subfields may lead to changes in representation by describing how memories are reactivated in the hippocampus itself. However, Molitor et al. (2021) instead describe how changes in hippocampal representations were related to the strength of reactivation in occipital, lateral parietal, and posterior cingulate cortex. How might reactivation in each of these cortical regions arise, and how might it then be related to subsequent changes in hippocampal representations?

Communication between the hippocampus and these cortical regions is bidirectional. Input from occipital cortex reaches the hippocampus in part via perirhinal and parahippocampal cortex, which both project to the hippocampus through connections with entorhinal cortex. Similarly, entorhinal cortex mediates the transfer of information from the hippocampus back to occipital cortex (Duvernoy et al., 2005). The hippocampus also sends input to the posterior cingulate and regions of parietal cortex by way of the thalamus; cingulate and parietal cortex, in turn, send reciprocal projections back to the hippocampus via the parahippocampal cortex (Duvernoy et al., 2005).

Functionally, these bidirectional anatomic connections support the flow of information between the hippocampus and these cortical regions, consistent with the fact that representations of retrieved memory content are widely distributed (Favila et al., 2020; Ritchey and Cooper, 2020). Extant models of episodic memory have posited that memory retrieval in the hippocampus may trigger cortical reactivation in at least two ways, as follows: (1) by driving the activation of memory representations in regions that were engaged during initial perception, such as occipital cortex (Teyler and DiScenna, 1986; Wheeler et al., 2000; Tompary et al., 2016); and (2) by increasing its coupling with the default mode network (DMN), which is frequently implicated in internally oriented attention during retrieval (Buckner et al., 2008; Ritchey and Cooper, 2020). Indeed, past work has shown that the content of retrieved memories is represented in occipital cortex (when memories contain visual information; Favila et al., 2020), as well as in key nodes of the DMN, such as lateral parietal (Kuhl and Chun, 2014) and posterior cingulate cortex (Ritchey and Cooper, 2020). Recent work also suggests that cortical reactivation is correlated with concurrent reactivation in the hippocampus, providing further evidence for a link between memory representations across the hippocampus and cortex (Pacheco Estefan et al., 2019). As such, in the present study, reactivation in occipital, parietal, and cingulate cortex may serve as an indirect readout of memory reactivation in the hippocampus.

Cortical activity can also feed back to the hippocampus, exerting an influence on mnemonic processing. Previous work has shown that targeted, noninvasive, electromagnetic stimulation of lateral parietal cortex led to increased connectivity with the hippocampus and improved performance on a hippocampus-dependent memory test (Wang et al., 2014). While, to our knowledge, the specific means by which reactivation-related cortical activity influences memory representations in the hippocampus remains largely unknown, here we consider how such impact could be exerted: we propose that cortical input (e.g., from parietal cortex) serves to shift attention toward retrieved hippocampal memory representations (Cabeza et al., 2008), ultimately biasing the hippocampus toward a state of heightened memory retrieval (Tarder-Stoll et al., 2020). This shift toward retrieval may render hippocampal memory representations more labile (as suggested by work on memory reconsolidation; Nader et al., 2000), opening an additional window for the modification of existing representations during the encoding of new, related information. Although speculative, this idea may serve as a useful starting point for future investigations.

Together, the bidirectional communication between the hippocampus and cortex, both anatomically and functionally, provides a working model of how cortical reactivation might relate to changes in hippocampal memory representations. The hippocampus reactivates memories with varying degrees of strength based on the intrinsic properties of different subfields. The hippocampus then triggers widespread cortical reactivation, which may feed back to the hippocampus and promote further modification of overlapping memories. This raises the possibility that the observed integration and differentiation in the hippocampus result from both subfield-specific differences in hippocampal reactivation and feedback from cortical reactivation.

Reconciling Findings with the Nonmonotonic Plasticity Hypothesis

As discussed above, Molitor et al. (2021) suggest that the NMPH provides one potential framework for understanding their findings: at differing levels of reactivation, there are concomitant changes in representations (Ritvo et al., 2019). However, one challenge in this interpretation is that evaluating the predictions of NMPH would require sampling from the full reactivation continuum (i.e., weak, moderate, and strong levels of reactivation; Wammes et al., 2021). Although Molitor et al. (2021) continuously varied the perceptual similarity of the linking B items, they only measured changes in the representations of A and C items associated with highly similar B shapes, potentially limiting the range of reactivation strengths that could be examined. While the authors did leverage natural variation in reactivation within this single similarity condition, they then used a median split to compare hippocampal representations across only two levels of cortical reactivation: strong versus weak. Consequently, Molitor et al. (2021) were unable to fully characterize the predicted relationship between reactivation and changes in memory representations, leaving open the possibility that their findings only reflect a limited window into this larger predicted pattern.

In conclusion, Molitor et al. (2021) show that cortical memory reactivation drives dissociable representations in distinct hippocampal subfields, adding to a growing body of work characterizing the relationship between memory reactivation and representation changes. Future studies should continue to investigate this relationship by relating a continuous measure of reactivation to representation changes, and by exploring the unique contributions of both hippocampal and cortical reactivation to the modification of hippocampal representations. Elucidating the conditions under which differentiation and integration occur will reveal insights into how memories of our daily interrelated experiences are organized in the brain, as well as how they guide behavior.

References

Aly M, Turk-Browne NB (2018) Flexible weighting of diverse inputs makes hippocampal function malleable. Neurosci Lett 680:13–22.

Bakker A, Kirwan CB, Miller M, Stark CEL (2008) Pattern separation in the human hippocampal CA3 and dentate gyrus. Science 319:1640–1642.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Cabeza R, Ciaramelli E, Olson IR, Moscovitch M (2008) The parietal cortex and episodic memory: an attentional account. Nat Rev Neurosci 9:613–625.

Chanales AJH, Oza A, Favila SE, Kuhl BA (2017) Overlap among spatial memories triggers repulsion of hippocampal representations. Curr Biol 27:2307–2317.e5.

Dimsdale-Zucker HR, Ritchey M, Ekstrom AD, Yonelinas AP, Ranganath C (2018) CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. Nat Commun 9:294.

Duncan KD, Schlichting ML (2018) Hippocampal representations as a function of time, subregion, and brain state. Neurobiol Learn Mem 153:40–56.

Duvernoy HM, Cattin F, Risold P-Y (2005) The human hippocampus: functional anatomy, vascularization and serial sections with MRI. Berlin: Springer.

Favila SE, Chanales AJH, Kuhl BA (2016) Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. Nat Commun 7:11066.

Favila SE, Lee H, Kuhl BA (2020) Transforming the concept of memory reactivation. Trends Neurosci 43:939–950.

Grande X, Berron D, Horner AJ, Bisby JA, Düzel E, Burgess N (2019) Holistic recollection via pattern completion involves hippocampal subfield CA3. J Neurosci 39:8100–8111.

Kesner RP, Rolls ET (2015) A computational theory of hippocampal function, and tests of the theory: new developments. Neurosci Biobehav Rev 48:92–147.

Kim G, Norman KA, Turk-Browne NB (2017) Neural differentiation of incorrectly predicted memories. J Neurosci 37:2022–2031.

Kriegeskorte N (2008) Representational similarity analysis – connecting the branches of systems neuroscience. Front Syst Neurosci 2:4.

Kuhl BA, Chun MM (2014) Successful remembering elicits event-specific activity patterns in lateral parietal cortex. J Neurosci 34:8051–8060.

Molitor RJ, Sherrill KR, Morton NW, Miller AA, Preston AR (2021) Memory reactivation during learning simultaneously promotes dentate gyrus/CA 2,3 pattern differentiation and ca 1 memory integration. J Neurosci 41:726–738.

Nader K, Schafe GE, Le Doux JE (2000) Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature 406:722–726.

O'Mara SM, Commins S, Anderson M, Gigg J (2001) The subiculum: a review of form, physiology and function. Prog Neurobiol 64:129–155.

O'Reilly RC, McClelland JL (1994) Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. Hippocampus 4:661–682.

Pacheco Estefan D, Sánchez-Fibla M, Duff A, Principe A, Rocamora R, Zhang H, Axmacher N, Verschure PFMJ (2019) Coordinated representational reinstatement in the human hippocampus and lateral temporal cortex during episodic memory retrieval. Nat Commun 10:2255.

Ritchey M, Cooper RA (2020) Deconstructing the posterior medial episodic network. Trends Cogn Sci 24:451–465.

Ritvo VJH, Turk-Browne NB, Norman KA (2019) Nonmonotonic plasticity: how memory retrieval drives learning. Trends Cogn Sci 23:726–742.

Schapiro AC, Turk-Browne NB, Botvinick MM, Norman KA (2017) Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. Philos Trans R Soc Lond B Biol Sci 372:20160049.

Schlichting ML, Preston AR (2015) Memory integration: neural mechanisms and implications for behavior. Curr Opin Behav Sci 1:1–8.

Schlichting ML, Mumford JA, Preston AR (2015) Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. Nat Commun 6:8151.

Tarder-Stoll H, Jayakumar M, Dimsdale-Zucker HR, Günseli E, Aly M (2020) Dynamic internal states shape memory retrieval. Neuropsychologia 138:107328.

Teyler TJ, DiScenna P (1986) The hippocampal memory indexing theory. Behav Neurosci 100:147–154.

Tompary A, Duncan K, Davachi L (2016) High-resolution investigation of memory-specific reinstatement in the hippocampus and perirhinal cortex. Hippocampus 26:995–1007.

Wammes JD, Norman KA, Turk-Browne NB (2021) Increasing stimulus similarity drives nonmonotonic representational change in hippocampus. bioRxiv. Advance online publication. Retrieved April 21, 2021. doi: 10.1101/2021.03.13.435275.

Wang JX, Rogers LM, Gross EZ, Ryals AJ, Dokucu ME, Brandstatt KL, Hermiller MS, Voss JL (2014) Memory enhancement: targeted enhancement of cortical-hippocampal brain networks and associative memory. Science 345:1054–1057.

Wheeler ME, Petersen SE, Buckner RL (2000) Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc Natl Acad Sci U S A 97:11125–11129.

Yassa MA, Stark CEL (2011) Pattern separation in the hippocampus. Trends Neurosci 34:515–525.