

Citation: Bavassi, L., & Fuentemilla, L. (2024). Segregation-to-integration transformation model of memory evolution. *Network Neuroscience*, 8(4), 1529–1544. https://doi.org/10.1162/netn_a_00415

DOI:
https://doi.org/10.1162/netn_a_00415

Received: 7 February 2024
Accepted: 22 August 2024

Competing Interests: The authors have declared that no competing interests exist.

Corresponding Author:
Luz Bavassi
luzbavassi@gmail.com

Handling Editor:
Michael Cole

Copyright: © 2024
Massachusetts Institute of Technology
Published under a Creative Commons
Attribution 4.0 International
(CC BY 4.0) license



The MIT Press

RESEARCH

Segregation-to-integration transformation model of memory evolution

Luz Bavassi^{1,2}  and Lluís Fuentemilla^{3,4,5} 

¹Laboratorio de Neurociencias de la Memoria, IFIByNE - UBA, CONICET, Buenos Aires, Argentina

²Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina

³Department of Cognition, Development and Education Psychology, University of Barcelona, Barcelona, Spain

⁴Institute of Neuroscience (UBNeuro), University of Barcelona, Barcelona, Spain

⁵Bellvitge Institute for Biomedical Research, Hospitalet de Llobregat, Barcelona, Spain

Keywords: Neural network, Modularity, Memory, Consolidation, Reactivation, Plasticity

ABSTRACT

Memories are thought to use coding schemes that dynamically adjust their representational structure to maximize both persistence and efficiency. However, the nature of these coding scheme adjustments and their impact on the temporal evolution of memory after initial encoding is unclear. Here, we introduce the Segregation-to-Integration Transformation (SIT) model, a network formalization that offers a unified account of how the representational structure of a memory is transformed over time. The SIT model asserts that memories initially adopt a highly modular or segregated network structure, functioning as an optimal storage buffer by balancing protection from disruptions and accommodating substantial information. Over time, a repeated combination of neural network reactivations involving activation spreading and synaptic plasticity transforms the initial modular structure into an integrated memory form, facilitating intercommunity spreading and fostering generalization. The SIT model identifies a nonlinear or inverted U-shaped function in memory evolution where memories are most susceptible to changing their representation. This time window, located early during the transformation, is a consequence of the memory's structural configuration, where the activation diffusion across the network is maximized.

AUTHOR SUMMARY

The Segregation-to-Integration Transformation (SIT) model provides a framework for memory transformation based on changes in the neural network properties. SIT posits that memories shift from highly modular to less modular network forms over time, driven by neural reactivations, activation spread, and plasticity rules. The SIT model identified a critical period, shortly after memory formation, where reactivations can induce significant structural modifications. As repeated reactivation passes, the network becomes more stable and integrated, becoming more resistant to change, thus preserving the core information while reducing the likelihood of distortion or loss.



The MIT Press

INTRODUCTION

Memories, like other natural systems, evolve to enhance survival in organisms. They adapt their representational structures to meet experiential needs better, aiming for optimized persistence and efficacy. While there is an agreement that memory evolution involves representational changes, the balance between persistence and efficacy is still not fully understood. A major research challenge is clarifying how experiences that enhance efficacy influence memory changes, ultimately promoting persistence.

Our current understanding of memory evolution suggests that, initially, memories are encoded by groups of neurons with synchronized activity, forming clustered neural ensembles that, if reactivated, induce memory retrieval (Josselyn & Tonegawa, 2020). These neural ensembles are identified in the hippocampus, but over time, they undergo structural changes, losing their modular properties in the hippocampus (Gonzalez, Zhang, Harutyunyan, & Lois, 2019), and other neural ensembles beyond the hippocampus gradually assume their representation (Frankland & Bontempi, 2005). This transition, however, occurs relatively slowly and requires repeated reactivation of the neural ensembles to shift the structure of the neural ensembles between networked regions (Frankland & Bontempi, 2005). For a while, the dominating view was that once memory neural ensembles shifted beyond the hippocampus, they became consolidated and stable in the long term. More contemporary views advocate that consolidation and reactivation may act in the service of generalization (Sun, Advani, Spruston, Saxe, & Fitzgerald, 2023). That is, given that individual memorized experiences rarely repeat exactly, generalization allows us to identify systematic relationships between features of the world, ultimately involving extending learned information to novel contexts. Thus, generalization involves linking and extracting commonalities among various memories. As the brain generalizes, initial memory representations transform their structure, becoming intricately connected to related memories (Nadel & Moscovitch, 1997; Winocur, Moscovitch, & Bontempi, 2010). Through this interconnected web of associations, the initially clustered memory representations become part of a broader network, where each memory influences and is influenced by others. These processes ensure that consolidated memories are not isolated but integrated into an interconnected form, making them applicable to various situations and contributing to the adaptive evolution of memories (Shohamy & Wagner, 2008; Sutherland & Lehmann, 2011). Thus, while most extant views of memory evolution offer insights into memory reorganization and generalization, developing a formal model that encapsulates the dynamic restructuring of memory networks remains an open area of research. Such a model would advance our understanding of memory consolidation, generalization, and the complex associations involved in adaptive memory evolution.

We propose the Segregation-to-Integration Transformation (SIT) model to account for how memory representations are transformed via repeated reactivations over time. SIT posits that memories are represented as neural network ensembles, with neurons (nodes) linked by connections (edges), and that memory transformation occurs through changes in the connectivity patterns between these ensembles (Ortega-de San Luis, Pezzoli, Urrieta, & Ryan, 2023; Ryan, Ortega-de San Luis, Pezzoli, & Sen, 2021; Ryan, Roy, Pignatelli, Arons, & Tonegawa, 2015; Tonegawa, Pignatelli, Roy, & Ryan, 2015). The SIT model assumes that memory networks are initially represented by information of single events encoded by groups of tightly connected neurons, sparsely linked with other neuronal groups, that represent other episodes with a common element or feature. This common element may be a concept, a temporal window, or a causal relationship (de Sousa, Chowdhury, & Silva, 2021). Highly modular networks serve as buffers to contain activation within their originating community rather than spreading to other

Modularity:

An index of the degree to which network nodes are organized into discrete communities.

communities (M. E. J. Newman, 2006). This modularity enhances storage capacity (Brunel, 2016) and represents an optimized configuration for early memories that balances the need to shield against disruptions while accommodating substantial information storage (Gonzalez et al., 2019). For example, the memory of a day at work may include memories of four different event experiences: working on a specific project, giving a lab talk, attending a lab meeting, and lecturing at the university (Figure 1A). While memories for these specific events can be stored and retrieved independently, after repeated encounters of similar experiences, they may become intricately linked into a broader memory structure that gives rise to an integrated form representing the experience of “my research activities.” We propose that, over time, repeated reactiverations transform the highly modular network structure into a more integrated form, incorporating links to nodes that represent overlapping information. This transition to a lower modular structure facilitates memory retrieval (Watrous, Tandon, Conner, Pieters, & Ekstrom, 2013; Westphal, Wang, & Rissman, 2017), intercommunity spreading, and promoting generalization. By investigating the dynamic coding schemes of memories, we seek to elucidate the mechanisms by which memories adapt to enhance survival and optimize functionality over time.

RESULTS

The SIT model seeks to analyze the evolution of memory within the context of network dynamics. SIT model focuses on the spatial organization of the network to represent memory accessibility, defined by how activation spreads through the network when specific nodes are activated. We propose that the initial memory state is represented by a set of nodes exhibiting a highly segregated structure that gradually shifts toward an integrated form induced by repeated reactiverations (Figure 1B). In a segregated network, activation is more likely to remain confined within specific clusters or communities, leading to a more context-specific and isolated recall. Conversely, in a highly integrated network, activation can spread more easily and efficiently across the network, facilitating the retrieval of interconnected memory components. This means that a memory cue can activate a broader, more cohesive set of related memories, enhancing holistic recall. The Degree of Integration, therefore, reflects the memory network’s structural configuration and its functional capacity to support different types of memory retrieval processes.

The SIT model posits that memory transformation is represented by changes in network structure, which depends on a shift from a segregated to an integrated form. Analytically, this can be assessed by quantifying the number of links between communities relative to the total number of possible links in the network. We labeled the output of this computation as the Degree of Integration (Z). $Z \sim 0$ implies a network strongly segregated, whereas $Z \sim 1$ yields more links between communities, that is, a more integrated structure (Figure 1B). Thus, we can study the segregation-to-integration dynamics over time by computing it for each resulting network after a reactivation.

We modeled memory reactivation following a deterministic rule that can induce memory network reorganization (Figure 1C). An external input or cue that overlaps with the original experience triggers memory reactivation, and a rate of nodes of the memory network is turned on (named as intensity, Int), transitioning its mode from an inactive to an active state (from $0 \rightarrow 1$). We assumed that a larger overlap between the current input and the existing memory would increase the reactivation’s intensity (Int) (Wammes, Norman, & Turk-Browne, 2022). The nodes turned on initialize the propagation of the activation over the network based on a Diffusion Linear Model (Nematzadeh, Ferrara, Flammini, & Ahn, 2014). Although

Segregated network:

A network that has a high number of communities.

Integrated network:

A network with a homogeneous connection distribution between its nodes.

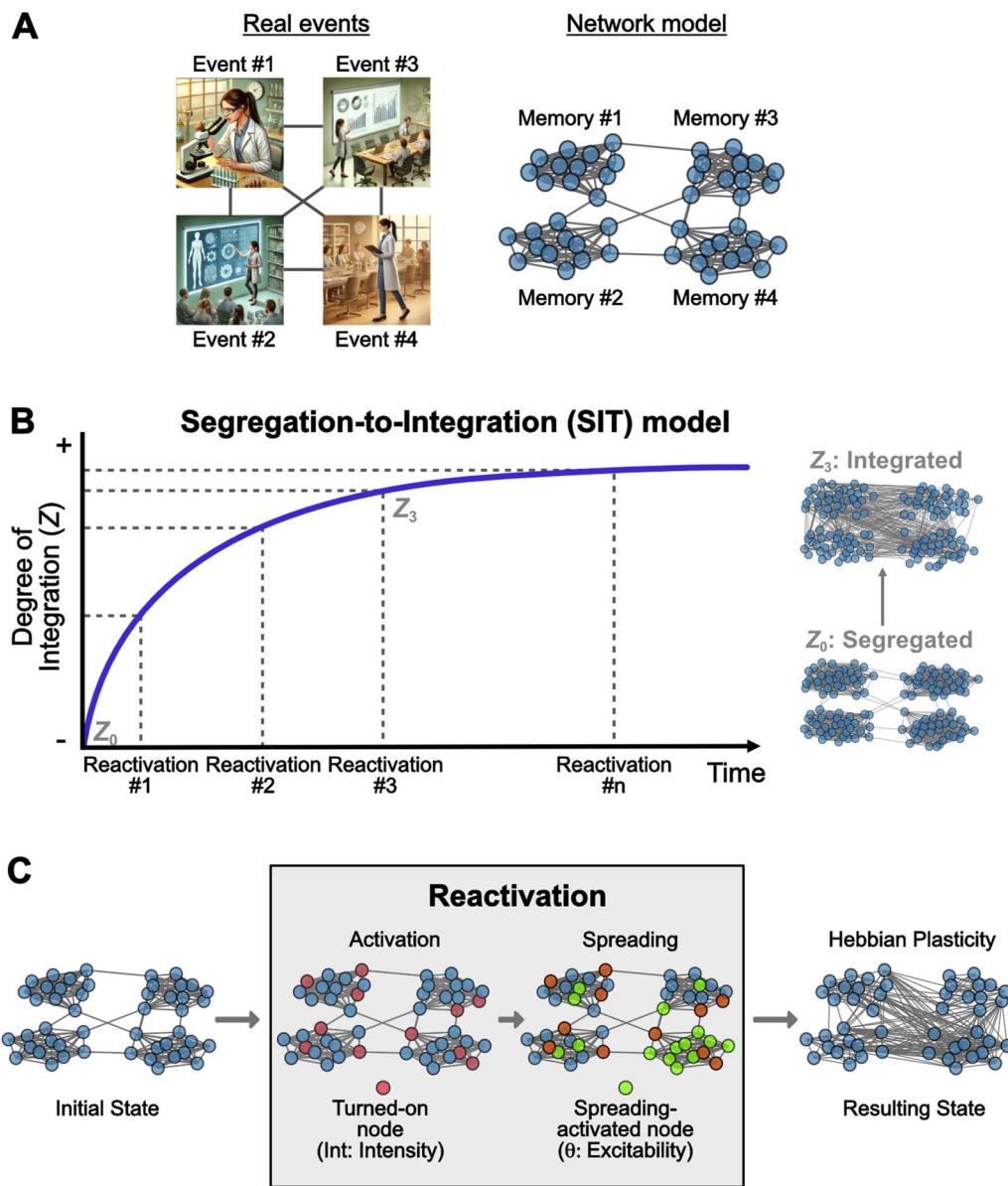


Figure 1. Schematic depiction of the SIT model. (A) The representation of memory as networks. Four separate experiences sharing a common element, for example, research activities, are represented by four subgroups of nodes tightly connected within them and sparsely linked with the nodes of the rest of the network. (B) The central tenet of the SIT model is that memory networks tend to shift from highly segregated to an integrated configuration guided by repeated reactivations over time. The Degree of Integration (Z) reflects the structural properties of the network, $Z \sim 0$ implies a network strongly segregated, whereas $Z \sim 1$ represents a network with a more integrated structure. (C) Memory reactivation induces retrieval and reorganization of the memory network. An external input that overlaps with the original experience triggers memory reactivation, and a rate of nodes of the memory network is turned on (Int), transitioning its mode to an active state. These turned-on nodes initialize the propagation of the activation over the network. The extent of the spreading is controlled by the excitability threshold parameter θ . Once the network achieves the steady state, a Hebbian plasticity rule modifies the connectivity pattern, creating and removing links between nodes.

it has been effective in capturing recurrent information transmission and retention in social networks (Centola, 2010; Danon, Arenas, & Díaz-Guilera, 2008), lately, it was proposed as a possible mechanism of brain network communication, both in the micro- and macroscopic scale. The Diffusion Linear Model included an excitability threshold parameter θ to control

the extent of the propagation. We conceived this threshold parameter as a proxy for the neuronal excitability state of network nodes. By systematically varying its value in our model, we assessed whether fluctuations in neuronal excitability influence memory ensemble interactions, promoting either integration or separation (Josselyn & Frankland, 2018). Finally, we implemented a plasticity rule that modifies the connectivity pattern, linking active nodes to promote pairwise activity, similar to Hebbian plasticity (Hebb, 2005). Additionally, we included a weakening rule that removes connections between active and inactive nodes (Stent, 1973).

To investigate memory transformation, we prepared an ensemble of highly segregated networks ($Z = 0.01$) with 128 nodes, divided equally into four communities of 32 nodes each. We then simulated reactivation by randomly activating 30% of each community's nodes ($Int = 0.3$) with a fixed excitability threshold at $\theta = 0.4$. We studied how each reactivation induced changes in the memory network by computing the Degree of Integration (Z). In line with the SIT hypothesis, the results of this analysis revealed that Z increased monotonically due to repeated reactivations, confirming that reactivating overlapping nodes effectively shifted the original segregated network memory configuration into an integrated network form in our model. We also found that this configural network change occurred rapidly within the first four reactivations, after which Z increased minimally with further reactivations (Figure 2A). These findings indicate that the mathematical rules used to simulate reactivation align with the central tenet of the SIT model: Memory undergoes a one-way transformation from a segregated form to an integrated configuration.

Network shifts from a segregated to an integration form should entail changes in how the spreading is distributed through the network along the process. Therefore, changes in the spatial structure of the network may be a determinant in modulating memory accessibility during its evolution. We examined this issue by computing network entropy (H). Entropy measures the information flow over a network based on the probability of jumping between nodes when traveling randomly through a network (Wiedermann, Donges, Kurths, & Donner, 2017). Sparse network configurations involve less entropy (i.e., $H \sim 0$) as there is low uncertainty along which path the information is transferred upon activation. In contrast, fully connected network configurations involve high entropy ($H \sim 1$), as any activated node has high uncertainty in the activation spread. In the context of the SIT model, we found that entropy was relatively high (i.e., $H > 0.5$) in all stages of network evolution (Figure 2B). However, the entropy curve peaked around the second reactivation, indicating the highest efficiency in information transmission at this point. This peak occurred before the network became highly integrated, as described in our previous analysis, thereby suggesting the existence of a critical time window during memory evolution at which a reactivation can access a larger portion of the stored memory.

The findings of maximum activation transfer at early stages of memory evolution indicate the existence of an optimal stage when network memory is most accessible as a whole. This suggests the existence of a specific stage in memory evolution where reactivation may induce greater changes in connectivity patterns, resulting in a higher degree of malleability. To account for this possibility, we computed the Malleability Index (ΔL ; Figure 2C), which quantifies changes in the number of connections within the network. The Malleability Index is expected to be low if a reactivation induces minimal changes to the existing network connectivity configuration, whereas it would be high if a reactivation elicits substantial changes to the existing number of connections. In the context of SIT, we expect low Malleability Index values when the memory network reaches a stable integrated form

Spatial structure of the network:
The pattern of network connectivity, which describes the configuration of the wiring diagram.

Malleability:
Changes in the number of connections or links of the network.

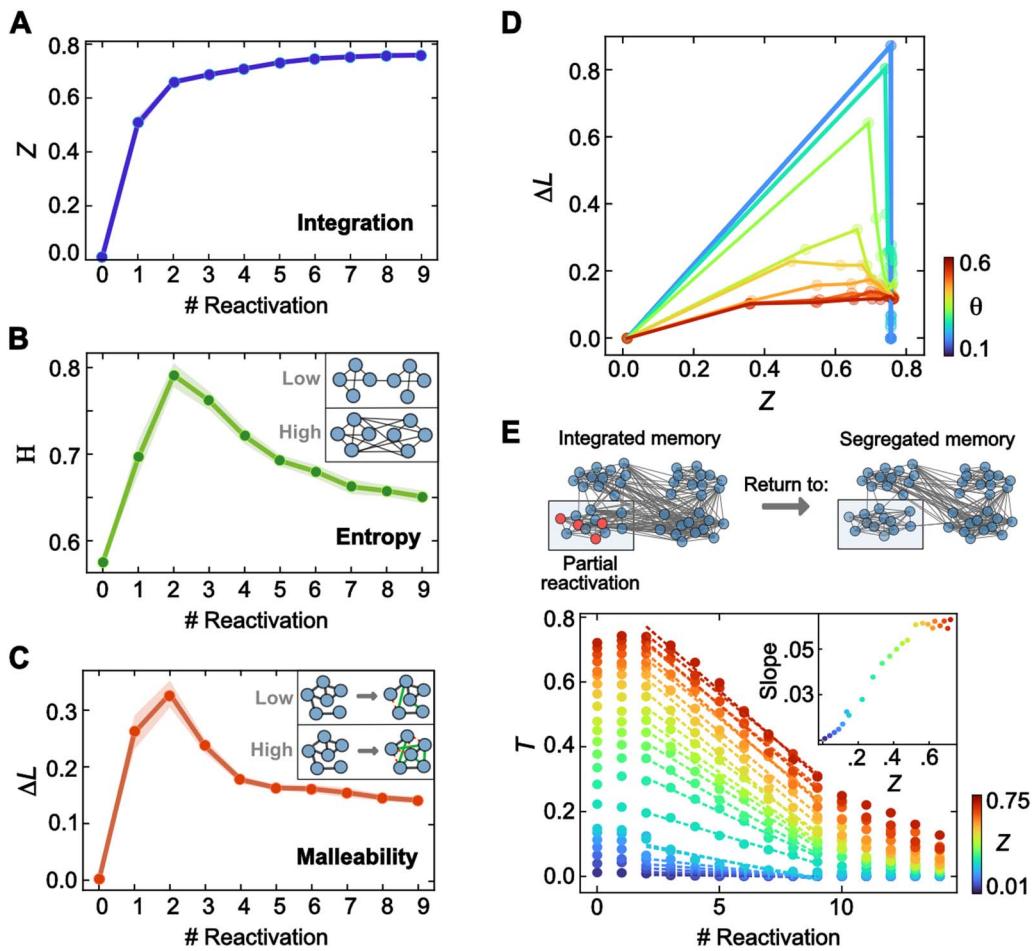


Figure 2. Model simulations. (A) Evolution of the Degree of Integration (Z) over 10 repeated reactivations. Z increases monotonically up to a highly integrated configuration. (B) Network entropy's evolution (H) over repeated reactivations. All the networks present high entropy levels, reaching the maximum early when the network is not completely integrated. (C) Malleability index (ΔL) over reactivations. ΔL presents a similar pattern to entropy. Intensity ($Int = 0.3$ per community), the excitability parameter ($\theta = 0.4$), and the initial structural configuration ($Z_0 = 0.01$) were fixed for the reported data in panels A, B, and C. (D) Relationship among Z and ΔL for different excitability thresholds for fixed intensity ($Int = 0.3$). The green trajectory is the conjunction of panels A and C. (E) The tightness index evolution over 15 partial repeated reactivations (only one community was reactivated with fixed intensity, $Int = 0.3$). Colors represent the initial network's Degree of Integration (Z). The excitability parameter was fixed at 0.4. The inset panel is the linear regression slope of tightness over reactivation numbers 3 and 9. The data reported are an average of multiple experimental runs (25) where the identity of the turned-on nodes is randomized in each run.

(i.e., during later reactivation stages). In contrast, we expect higher Malleability Index values at the early stages of memory transformation, aligning with the previously observed entropy function. This analysis revealed that the ΔL values evolved nonlinearly throughout memory transformation, displaying an inverted U-shaped function. Initial reactivations induced substantial changes in the connectivity pattern of the memory networks, reaching a maximum ΔL at around four to five reactivations. However, after achieving the maximum, ΔL decreased and flattened around a low level above 0. This suggests the existence of a specific stage in memory evolution where reactivation may induce greater changes in the wire patterns, resulting in a higher degree of malleability, and implies that once this malleability window has passed, reactivation still induces memory changes, albeit at a minimal level.

The low-dimensional reactivation proposed in the SIT model involves a unique parameter, the excitability threshold (θ). It is an intrinsic parameter of spreading, as it limits the degree of activation throughout the network. We explored the extent to which the SIT findings depended on variation on this parameter in a network by calculating the trajectories of memory evolution in the $Z - \Delta L$ space for different excitability thresholds θ (Figure 2D). The nonlinear relation between Z and ΔL is displayed as an open loop in the phase space (green line; $\theta = 0.4$). As expected, the amplitude of the nonlinearity increases for more permissive thresholds. That is, lower excitability thresholds favored the spreading of activation. After the maximum malleability occurred, memories involving more permissive excitability thresholds did not present changes in the number of connections (lack of malleability). In late reactivations, memories with stringent thresholds showed limited variability in the Malleability Index. Overall, the excitability threshold modulated the memory's malleability during the early phase of the process and constrained the potential for change in an integrated memory.

While the results of the SIT model revealed a one-way tendency of network memories to shift from a segregated to an integrated form, sometimes, preserving a discrete representation of memory over time is necessary. Theoretical models have shown that this can be achieved through pattern separation mechanisms that encode an orthogonal representation of the original memory (Favila, Chanales, & Kuhl, 2016; O'Reilly & Norman, 2002; Treves & Rolls, 1994) or by reinforcing the memory individually through repeated reactivation (Lohnas et al., 2018). We, therefore, asked whether a network's community can remain discrete and identifiable within the network upon its reactivation. To address this, we implemented the repeated reactivation of 30% of nodes from one original memory (one community), even at states where the network structure has already reached a relatively integrated form (Figure 2E). We defined the tightness index (T) as a measure of the strength of the community (the number of links outside the community over the total number of links). $T \sim 0$ implies almost null links outside the community, implying an isolated community; $T \sim \frac{1}{2}$ implies the same number of links within the community as to the rest of the network, while a community with a small number of internal links achieves $T \sim 1$. This served as a proxy to determine how well the memory returned to a segregated form or how it resisted integration. This analysis revealed that the T value decreased for all Z states of the original network, reaching a low value (i.e., a highly segregated network) after several reactivations. These findings indicate that even when the original network was in highly integrated states, the repeated reactivation of partial elements from one of the four communities could return the network to a segregated form, irrespective of its original state of integration. However, when we measured the slope of the decreasing T function, we found that the speed of this returning function varied nonmonotonically throughout the Z state. Specifically, T slowed its rate of decrease linearly from low Z to intermediate Z states (approximately 0.6), upon which it reached a similar decrease function (inset Figure 2E). The model shows that reactivation requires elements of different communities to achieve the integration process. If reactivation involves elements of only one community, the memory becomes isolated, remaining protected.

So far, our findings have been based on a fixed network size (i.e., $N = 128$) and fixed reactivation strength (i.e., $Int = 0.3$). While these parameters are set arbitrarily in our modeling, they may influence the temporal dynamics of network transformation. Therefore, an open question that remains to be explored is whether our results can be generalized when these parameters are varied across their full range of values.

First, we tested that the memory transformation described did not depend on the size of the network by expanding our analysis to include networks of various sizes (for fixed $Int = 0.3$ and $\theta = 0.4$). We analyzed the dynamics of the Degree of Integration (Z) and the Malleability Index

(ΔL) for highly segregated networks consisting of 16, 32, 64, 128, 256, 512, and 1,024 nodes (Figure 3A and B). Each network was divided into four equal communities with different initial Degree of Integration ($Z_{16} = 0.3$; $Z_{32} = 0.09$; $Z_{64} = 0.02$; $Z_{128} = 0.01$; $Z_{256} = 0.001$; $Z_{512} = 0.0004$; $Z_{1024} = 0.0001$). The Degree of Integration (Z) exhibited similar behavior across memory networks of different sizes, showing a monotonic increase in Z during reactivations and reaching a fully integrated state early on. The speed of integration did not appear to depend on the network size, although the smallest networks achieved a more integrated configuration. However, the curves of malleability displayed different profiles for networks of different sizes. The inverted U-shaped function was observed only in small networks. In contrast, larger networks exhibited a monotonic increase in the malleability function, probably

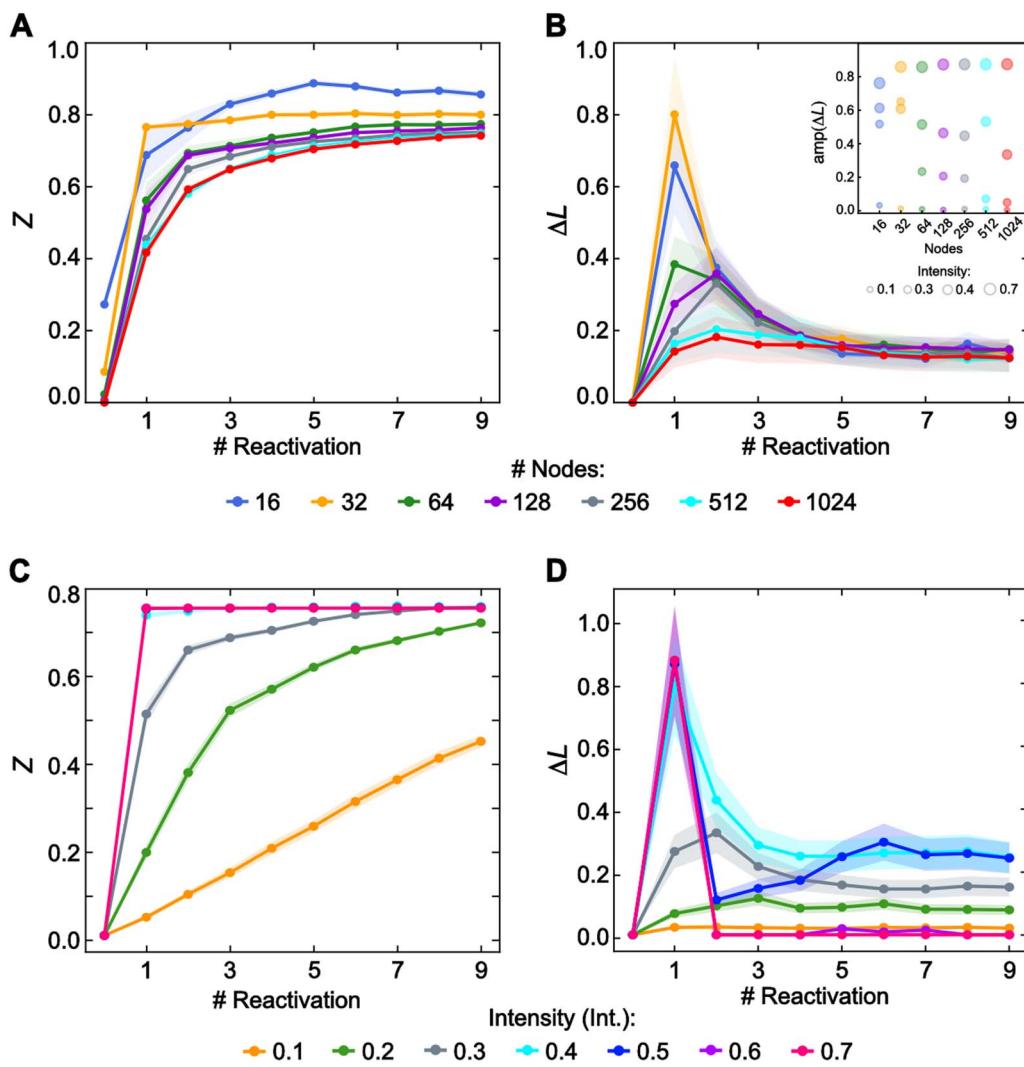


Figure 3. Model generalization. (A) Degree of Integration and (B) Malleability Index over 10 repeated reactivations for different network's sizes ($N = 16, 32, 128, 256, 512, 1024$). Intensity ($Int = 0.3$) and excitability ($\theta = 0.4$) were fixed. The inset panel in B depicted the maximum of malleability for different network sizes, calculated by subtracting the peak value of the function from the averaged value of the five last values. The size of the circles represented the intensity level ($\theta = 0.4$). (C) Degree of Integration and (D) Malleability Index over 10 repeated reactivations for different intensities. The networks had 128 nodes and four communities; the excitability parameter was fixed at $\theta = 0.4$. The data reported are an average of multiple experimental runs (25) where the identity of the turned-on nodes is randomized in each run.

because the nonlinearity rose with a different set of parameters (more permissive excitability thresholds or higher intensity values). To disentangle this, we depicted the maximum amplitude of malleability for different network sizes and intensity levels (with a fixed excitability threshold $\theta = 0.4$; inset Figure 3B). As expected, the U-shaped function appeared for all network sizes, but larger networks need higher intensities to display the nonlinearity. Interestingly, all the networks showed the same level of malleability after achieving their maximum.

Finally, we analyzed the role of intensity in our findings. Intensity (Int) refers to the rate at which network nodes are activated in each reactivation. Figure 3C and D display the Degree of Integration (Z) and Malleability Index (ΔL) for a network of four communities (32 nodes per community) with a fixed excitability threshold ($\theta = 0.4$). Starting with a segregated configuration ($Z = 0.01$), networks reactivated with the highest intensities achieved a highly integrated state after the first activation. In contrast, the lowest intensities induced a slow integration process, with networks reactivated at $Int = 0.1$, failing to achieve an integrated configuration even after 10 reactivations, requiring additional activations to reach the final state. The Malleability Index (ΔL) exhibited a U-shaped function for all $Int > 0.2$. Notably, networks reactivated with the highest intensities showed no changes in malleability after reaching the maximum.

In summary, the SIT model is robust, extending to memories of different sizes and reactivations with varying intensity strengths.

DISCUSSION

The SIT model provides a framework for understanding how memory evolves, summarized through changes in neural network properties. SIT posits that memories shift from highly modular to less modular network forms over time, driven by neural reactivations, activation spread, and plasticity rules. The SIT model identified an optimal window of memory malleability, indicating a nonlinear, inverted U-shaped function, where memories are most malleable early on and gradually become less changeable. This suggests the existence of a critical period shortly after memory formation, during which the memory network is highly plastic and susceptible to changes. During this period, reactivations can induce significant structural modifications, allowing for the integration of new information and the strengthening or weakening of existing connections. As time progresses and the memory undergoes repeated reactivations, the network becomes more stable and integrated, leading to a decrease in malleability. This stabilization process implies that the memory is becoming more resistant to change, thus preserving the core information while reducing the likelihood of distortion or loss.

An underlying assumption of the SIT model is that the transformation of memory relies on modifications in connectivity patterns among nodes within the memory network. This perspective aligns with the idea that learning involves modifications to the wiring diagram of a neural ensemble, where previously unconnected units establish connections and vice versa (Ortega-de San Luis et al., 2023; Ryan et al., 2015, 2021; Tonegawa et al., 2015). However, while alterations in the wiring pattern offer a potential substrate for encoding more extensive information in sparse coding models (Chklovskii, Mel, & Svoboda, 2004; Knoblauch, Palm, & Sommer, 2010), it also underscores that other forms of plasticity based on changes in synaptic weights, the strength of connections between cells, are crucial for comprehending memory evolution (Bonhoeffer & Yuste, 2002; Chklovskii et al., 2004). Although both mechanisms are likely involved in engram cell formation and function (Poo et al., 2016), investigations of network

properties in hippocampal code representation have shown that wiring diagrams effectively encode specific experiences (Ortega-de San Luis et al., 2023; Ryan et al., 2021). In addition, shifts in these connectivity patterns provide a more robust explanation for memory transformation than alterations in the individual firing properties of the neurons (Gava et al., 2021). In subsequent work, the simplicity of the model may enable us to readily incorporate new variables that account for changes in synaptic weights encompassing both modifications in the wiring diagram and the strength of the connections.

The SIT model identified an optimal window of memory malleability, where memories are most malleable early on and gradually become less changeable. This result may reconcile theoretical models of memory consolidation (McClelland, McNaughton, & O'Reilly, 1995; Nadel & Moscovitch, 1997; Squire, 2004; Winocur & Moscovitch, 2011) and reconsolidation (Nader & Hardt, 2009; Sara, 2000), which both suggest that memories change over time. However, while reconsolidation posits one-shot changes, consolidation highlights continuous evolution (Dudai & Eisenberg, 2004). Our model shows that memories can evolve perpetually but have an optimal window of malleability early on. This aligns with the idea that young memories are more susceptible to disruption, while older ones are more stable (Alberini, 2011; Eisenberg & Dudai, 2004; Frankland & Bontempi, 2005; Milekic & Alberini, 2002; Suzuki et al., 2004). After this window, memories can still change, but minimally (McKenzie & Eichenbaum, 2011; Roüast & Schönauer, 2023).

We conceptualized that the network's excitability threshold and the reactivation's intensity modulated structural changes in the memory network. The excitability threshold was included to reflect findings from rodent studies showing that memory allocation depends on the intrinsic excitability of cells (Josselyn & Frankland, 2018). Fluctuations in neuronal excitability dictate how engrams interact, promoting either memory integration or separation. The intensity of reactivation manipulated neuron coactivation based on the overlap between the current input and existing memory (Wammes et al., 2022). Recent theoretical models, like the nonmonotonic plasticity hypothesis, suggest a U-shaped pattern of representational change: Low coactivation results in no change, high coactivation strengthens connections, and moderate coactivation leads to differentiation (Detre, Natarajan, Gershman, & Norman, 2013; E. L. Newman & Norman, 2010; Ritvo, Turk-Browne, & Norman, 2019). Our model found that excitability and intensity affected the rate of memory transition from segregated to integrated forms depending on the network size. Higher excitability and intensity increased propagation and memory malleability, while lower levels restricted activation spread and decreased malleability. Intermediate levels produced distinct malleability peaks, with maximum malleability occurring when the memory configuration was neither very segregated nor fully integrated, aligning with the nonmonotonic plasticity hypothesis.

The SIT model proposes that the shift of memories from segregated to integrated forms responds to the adaptation of the memory system to highly dynamic environments, where novel and past experiences may only partially overlap, necessitating a balanced interplay between stored information and accessibility. In this context, highly segregated memories may function suboptimally, owing to their limited accessibility (Hintze & Adami, 2008). In contrast, highly integrated memory network forms prioritize accessibility over storage quality, offering an adaptive structural configuration in ever-changing contexts. Network science has shown that at the intersection of these two extreme network configurations, there exists an optimal modularity configuration that effectively balances memory storage and information diffusion (Nematzadeh et al., 2014; Rodriguez, Izquierdo, & Ahn, 2019). Similarly, we propose that such an optimal modularity structural state of a memory form is ideal for inducing changes

because it provides a balanced trade-off between the probability of being reactivated and the diffusion of activation throughout the network.

While the SIT model provides a low-dimensional framework in which memory is proactively shaped into a generalized form that involves an integrated state, it fails to address important issues. One such open question pertains to the distribution of network configuration and how the network configuration and its transformation are distributed in the brain. Another remaining issue would be to study the effects of memory reactivation under different brain states, such as wake and sleep stages, that are thought to determine the nature of memory transformation. For example, reduced excitability and synaptic downscaling are thought to occur during the Slow Wave Sleep stage, but increased reactivation intensity and synaptic potentiation may better take place during the rapid eye movement (REM) sleep stage. While we await a better characterization of these issues, our model offers a simple guiding principle based on the topological structure of neural networks by which memories transform over time.

MATERIALS AND METHODS

Simulations were done using homemade Python codes, mainly using Networkx (<https://networkx.org/>; a package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks).

Building Networks

The Segregated-to-Integrated Model studied the evolution of a set of one-layer, unweighted, and undirected networks. Each node could achieve one of two possible values: $s = \{0, 1\}$, where 1 represented the “active” mode and 0 was the “inactive” one.

Initially ($T = 0$), the networks have $N = \{16, 32, 128, 256, 512, 1024\}$ nodes split into $C = 4$ equal communities of N/C nodes each. Each community is a random regular graph, where each node has the same number of neighbors. We then linked the communities with L^{ext} edges assigned randomly to pairs of nodes in different communities. This arrangement ensured that every node and community was not isolated and that they played a similar role in the network.

The mean node degree within the community was set to $k_{\text{int}} = \frac{1}{2}N/C$ and the number of edges that connected different communities varies from 1 to $N - 1$. To formalize the state of the network, we quantified the Degree of Integration following Equation 1:

$$Z = \frac{1}{L} \sum_i^C \sum_{\substack{n \in \mathcal{C}_i \\ m \notin \mathcal{C}_i}} L_{n,m}^{\text{ext}} \quad (1)$$

where L is the total number of edges and $L_{n,m}^{\text{ext}}$ are edges that linked node n , a member of the initial community \mathcal{C}_i , with node m that belongs to another community. Thus, $Z \sim 0$ defines a highly segregated network, while $Z \sim 1$ implies a network fully integrated (Figure 1B).

Reactivation

Computationally, memory reactivation involves a sequence of three deterministic rules: First, we activate a random set of nodes (turn-on); second, we spread the activation throughout the network (spreading); and finally, we modify the connection pattern of the network (plasticity).

Turn-on. The reactivation intensity (*Int*) is the fraction of inactive nodes that transition to the active state (from 0 → 1). It ranged from 0.1 to 0.7, with increments of 0.02. To introduce variability, the number of activated nodes per community was determined by sampling from a normal distribution with a mean of *Int* and a standard deviation of 0.05. To decide which nodes of the community become active, we used a uniform random number generator that selects values between 1 and *N/C*.

Spreading. We adopted a Diffusion Linear Threshold model (Nematzadeh et al., 2014) to achieve activity propagation over the network. The state of node *n* is updated according to the following linear threshold rule, Equation 2:

$$S_n(t+1) = \begin{cases} 1 & \text{if } \theta k_n < \sum_{j \in N_n} S_j(t) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where θ is the excitability threshold parameter, k_n was the node degree, and $j \in N_n$ represented the neighbors of node *n*. Throughout the simulations, θ varied between 0.1 and 0.6, with increments of 0.02. Once a node becomes active, it remains in this state. The spreading persisted until the network reaches a stable state, defined as $S_n(t+1) = S_n(t)$, for all *n* or exceeds 50 iterations (Nematzadeh et al., 2014).

Plasticity. Once the network reaches a stable state (*T*), we implement a plasticity rule. This step modifies the connectivity matrix of the neural network by creating/removing an edge between two nodes $L_{n,m}$ at *T* + 1. If both nodes are in an active state, they are connected, and the link between both is removed if the nodes are uncorrelated (one in an active mode and the other in an inactive state; Equation 3):

$$L_{n,m}(T+1) = \begin{cases} 1 & \text{if } S_n(T) = S_m(T) = 1 \\ 0 & \text{if } S_n(T) \neq S_m(T) \\ L_{n,m}(T) & \text{otherwise} \end{cases} \quad (3)$$

Memory Transformation

Memory transformation is simulated by applying a sequence of reactivations (10 or 15). The intensity (*Int*) remains constant throughout each reactivation. The nodes activated in each iteration were randomly assigned, as described above. We repeated the entire network transformation with the same set of parameters 25 times to characterize the mean behavior independently of the random assignment of nodes.

Entropy

Network entropy is based on the classical Shannon entropy for discrete distributions (Freitas, Aquino, Ramos, Frery, & Rosso, 2019). It relies on the probability that a random walker goes from node *n* to any node *m*. The probability distribution $P^{(n)}$ has entries $p_{n \rightarrow m} = \frac{1}{k_n}$ (if *n* is connected with *m*) and 0 otherwise. So, the entropy per node is defined (Equation 4):

$$h_n(P^{(n)}) = -\sum_{m=1}^{N-1} p_{n \rightarrow m} \ln(p_{n \rightarrow m}) = \ln(k_n) \quad (4)$$

and $h_n(P^{(n)}) = 0$ if node *n* is fully disconnected.

Finally, the average normalized entropy, called the network entropy (H), is as follows (Equation 5):

$$H(P) = \frac{1}{N \ln(N-1)} \sum_{n=1}^N h(P^{(n)}) \quad (5)$$

Malleability Index

To focus on the plasticity of memory, we study changes in the number of edges. The Malleability Index (ΔL) is the number of created (L^C) and removed (L^R) edges over the total number of edges before reactivation (L_{-1} ; Equation 6).

$$\Delta L = \frac{L^C + L^R}{L_{-1}} \quad (6)$$

ΔL is a positive parameter, and 0 value means no change in the number of edges.

To determine the existence of the inverted U-shaped function for networks of different sizes, we analyze the amplitude of the nonlinearity (substrating the final level of malleability) following Equation 7:

$$amp(\Delta L) = \max\{\Delta L\} - \langle \Delta L \rangle_{-5} \quad (7)$$

where $\langle \Delta L \rangle_{-5}$ is the mean of malleability of the last five reactivations.

Preserved Identity of a Community

To study how well a memory returned to a segregated form or how it resisted integration, we ran a sequence of 15 repeated reactivations for various networks of different Degree of Integration (Z between 0.001 and 0.6). The excitability threshold was $\theta = 0.4$, and the networks had 128 nodes. Reactivation intensity was fixed as $Int = 0.3$, and it only involved nodes of one community (\mathcal{C}). We repeated each network transformation with the same set of parameters 25 times. For each reactivation, we calculated the strength or cohesion of the reactivated community (\mathcal{C}) with the tightness measure (T_C ; Equation 8):

$$T_C = \frac{1}{L_C} \sum_{\substack{i \in \mathcal{C} \\ j \notin \mathcal{C}}} L_{i,j}^{ext} \quad (8)$$

where L_C is the total number of edges of the community C and $L_{i,j}^{ext}$ are edges that linked node i (a member of the community \mathcal{C}) with node j of the rest of the network. Thus, $T_C \sim 0$ defines a highly cohesive community while $T_C \sim 1$ implies a weak community, fully integrated with the rest of the network.

We analyze the slowed rate of T over the reactivations. We linearly fit the T 's evolution from reactivation number 3 to 9 and calculate the slope.

ACKNOWLEDGMENTS

This work was supported by a grant from the Agencia Nacional de Promoción Científica y Tecnológica PICT 2020 - 00956 to L.B. and from the Spanish Ministerio de Ciencia, Innovación y Universidades, which is part of Agencia Estatal de Investigación (AEI), through the project PID2022 - 140426NB - I00 (funded by MCIN/AEI/10.13039/501100011033/ and FEDER a way to make Europe) to L.F. We thank CERCA Programme/Generalitat de Catalunya for institutional support.

AUTHOR CONTRIBUTIONS

Luz Bavassi: Conceptualization; Data curation; Formal analysis; Methodology; Software; Visualization; Writing – original draft; Writing – review & editing. Lluis Fuentemilla: Conceptualization; Data curation; Investigation; Methodology; Project administration; Validation; Visualization; Writing – original draft; Writing – review & editing.

FUNDING INFORMATION

Luz Bavassi, Fondo para la Investigación Científica y Tecnológica (<https://dx.doi.org/10.13039/501100006668>), Award ID: PICT 2020 - 00956. This work was supported by the Spanish Ministerio de Ciencia, Innovación y Universidades, which is part of Agencia Estatal de Investigación (AEI), through the projects PID2019-111199GB-I00 and PID2022-140426NB-I00 to L.F. (funded by MCIN/AEI/10.13039/501100011033 and FEDER a way to make Europe).

REFERENCES

Alberini, C. M. (2011). The role of reconsolidation and the dynamic process of long-term memory formation and storage. *Frontiers in Behavioral Neuroscience*, 5, 12. <https://doi.org/10.3389/fnbeh.2011.00012>, PubMed: 21436877

Bonhoeffer, T., & Yuste, R. (2002). Spine motility: Phenomenology, mechanisms, and function. *Neuron*, 35(6), 1019–1027. [https://doi.org/10.1016/S0896-6273\(02\)00906-6](https://doi.org/10.1016/S0896-6273(02)00906-6), PubMed: 12354393

Brunel, N. (2016). Is cortical connectivity optimized for storing information? *Nature Neuroscience*, 19(5), 749–755. <https://doi.org/10.1038/nn.4286>, PubMed: 27065365

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329(5996), 1194–1197. <https://doi.org/10.1126/science.1185231>, PubMed: 20813952

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature*, 431(7010), 782–788. <https://doi.org/10.1038/nature03012>, PubMed: 15483599

Danon, L., Arenas, A., & Díaz-Guilera, A. (2008). Impact of community structure on information transfer. *Physical Review E*, 77(3), 036103. <https://doi.org/10.1103/PhysRevE.77.036103>, PubMed: 18517457

de Sousa, A. F., Chowdhury, A., & Silva, A. J. (2021). Dimensions and mechanisms of memory organization. *Neuron*, 109(17), 2649–2662. <https://doi.org/10.1016/j.neuron.2021.06.014>, PubMed: 34242564

Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, 51(12), 2371–2388. <https://doi.org/10.1016/j.neuropsychologia.2013.02.017>, PubMed: 23499722

Dudai, Y., & Eisenberg, M. (2004). Rites of passage of the engram: Reconsolidation and the lingering consolidation hypothesis. *Neuron*, 44(1), 93–100. <https://doi.org/10.1016/j.neuron.2004.09.003>, PubMed: 15450162

Eisenberg, M., & Dudai, Y. (2004). Reconsolidation of fresh, remote, and extinguished fear memory in Medaka: Old fears don't die. *European Journal of Neuroscience*, 20(12), 3397–3403. <https://doi.org/10.1111/j.1460-9568.2004.03818.x>, PubMed: 15610172

Favila, S. E., Chanales, A. J. H., & Kuhl, B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nature Communications*, 7(1), 11066. <https://doi.org/10.1038/ncomms11066>, PubMed: 27925613

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, 6(2), 119–130. <https://doi.org/10.1038/nrn1607>, PubMed: 15685217

Freitas, C. G. S., Aquino, A. L. L., Ramos, H. S., Frery, A. C., & Rosso, O. A. (2019). A detailed characterization of complex networks using Information Theory. *Scientific Reports*, 9(1), 16689. <https://doi.org/10.1038/s41598-019-53167-5>, PubMed: 31723172

Gava, G. P., McHugh, S. B., Lefèvre, L., Lopes-dos-Santos, V., Trouche, S., El-Gaby, M., ... Dupret, D. (2021). Integrating new memories into the hippocampal network activity space. *Nature Neuroscience*, 24(3), 326–330. <https://doi.org/10.1038/s41593-021-00804-w>, PubMed: 33603228

Gonzalez, W. G., Zhang, H., Harutyunyan, A., & Lois, C. (2019). Persistence of neuronal representations through time and damage in the hippocampus. *Science*, 365(6455), 821–825. <https://doi.org/10.1126/science.aav9199>, PubMed: 31439798

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory*. Psychology Press. <https://doi.org/10.4324/9781410612403>

Hintze, A., & Adami, C. (2008). Evolution of complex modular biological networks. *PLoS Computational Biology*, 4(2), e23. <https://doi.org/10.1371/journal.pcbi.0040023>, PubMed: 18266463

Josselyn, S. A., & Frankland, P. W. (2018). Memory allocation: Mechanisms and function. *Annual Review of Neuroscience*, 41, 389–413. <https://doi.org/10.1146/annurev-neuro-080317-061956>, PubMed: 29709212

Josselyn, S. A., & Tonegawa, S. (2020). Memory engrams: Recalling the past and imagining the future. *Science*, 367(6473), eaaw4325. <https://doi.org/10.1126/science.aaw4325>, PubMed: 31896692

Knoblauch, A., Palm, G., & Sommer, F. T. (2010). Memory capacities for synaptic and structural plasticity. *Neural Computation*,

22(2), 289–341. <https://doi.org/10.1162/neco.2009.08-07-588>, PubMed: 19925281

Lohnas, L. J., Duncan, K., Doyle, W. K., Thesen, T., Devinsky, O., & Davachi, L. (2018). Time-resolved neural reinstatement and pattern separation during memory decisions in human hippocampus. *Proceedings of the National Academy of Sciences*, 115(31), E7418–E7427. <https://doi.org/10.1073/pnas.1717088115>, PubMed: 30006465

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (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. *Psychological Review*, 102(3), 419–457. <https://doi.org/10.1037/0033-295X.102.3.419>, PubMed: 7624455

McKenzie, S., & Eichenbaum, H. (2011). Consolidation and reconsolidation: Two lives of memories? *Neuron*, 71(2), 224–233. <https://doi.org/10.1016/j.neuron.2011.06.037>, PubMed: 21791282

Milekic, M. H., & Alberini, C. M. (2002). Temporally graded requirement for protein synthesis following memory reactivation. *Neuron*, 36(3), 521–525. [https://doi.org/10.1016/S0896-6273\(02\)00976-5](https://doi.org/10.1016/S0896-6273(02)00976-5), PubMed: 12408853

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, 7(2), 217–227. [https://doi.org/10.1016/S0959-4388\(97\)80010-4](https://doi.org/10.1016/S0959-4388(97)80010-4), PubMed: 9142752

Nader, K., & Hardt, O. (2009). A single standard for memory: The case for reconsolidation. *Nature Reviews Neuroscience*, 10(3), 224–234. <https://doi.org/10.1038/nrn2590>, PubMed: 19229241

Nematzadeh, A., Ferrara, E., Flammini, A., & Ahn, Y.-Y. (2014). Optimal network modularity for information diffusion. *Physical Review Letters*, 113(8), 088701. <https://doi.org/10.1103/PhysRevLett.113.088701>, PubMed: 25192129

Newman, E. L., & Norman, K. A. (2010). Moderate excitation leads to weakening of perceptual representations. *Cerebral Cortex*, 20(11), 2760–2770. <https://doi.org/10.1093/cercor/bhq021>, PubMed: 20181622

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23), 8577–8582. <https://doi.org/10.1073/pnas.0601602103>, PubMed: 16723398

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, 6(12), 505–510. [https://doi.org/10.1016/S1364-6613\(02\)02005-3](https://doi.org/10.1016/S1364-6613(02)02005-3), PubMed: 12475710

Ortega-de San Luis, C., Pezzoli, M., Urrieta, E., & Ryan, T. J. (2023). Engram cell connectivity as a mechanism for information encoding and memory function. *Current Biology*, 33(24), 5368–5380. <https://doi.org/10.1016/j.cub.2023.10.074>, PubMed: 37992719

Poo, M.-M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., ... Stevens, C. (2016). What is memory? The present state of the engram. *BMC Biology*, 14, 40. <https://doi.org/10.1186/s12915-016-0261-6>, PubMed: 27197636

Ritvo, V. J. H., Turk-Browne, N. B., & Norman, K. A. (2019). Non-monotonic plasticity: How memory retrieval drives learning. *Trends in Cognitive Sciences*, 23(9), 726–742. <https://doi.org/10.1016/j.tics.2019.06.007>, PubMed: 31358438

Rodriguez, N., Izquierdo, E., & Ahn, Y.-Y. (2019). Optimal modularity and memory capacity of neural reservoirs. *Network Neuroscience*, 3(2), 551–566. https://doi.org/10.1162/netn_a_00082, PubMed: 31089484

Roüast, N. M., & Schönauer, M. (2023). Continuously changing memories: A framework for proactive and non-linear consolidation. *Trends in Neurosciences*, 46(1), 8–19. <https://doi.org/10.1016/j.tins.2022.10.013>, PubMed: 36428193

Ryan, T. J., Ortega-de San Luis, C., Pezzoli, M., & Sen, S. (2021). Engram cell connectivity: An evolving substrate for information storage. *Current Opinion in Neurobiology*, 67, 215–225. <https://doi.org/10.1016/j.conb.2021.01.006>, PubMed: 33812274

Ryan, T. J., Roy, D. S., Pignatelli, M., Arons, A., & Tonegawa, S. (2015). Engram cells retain memory under retrograde amnesia. *Science*, 348(6238), 1007–1013. <https://doi.org/10.1126/science.aaa5542>, PubMed: 26023136

Sara, S. J. (2000). Retrieval and reconsolidation: Toward a neurobiology of remembering. *Learning & Memory*, 7(2), 73–84. <https://doi.org/10.1101/lm.7.2.73>, PubMed: 10753974

Seguin, C., Sporns, O., & Zalesky, A. (2023). Brain network communication: Concepts, models and applications. *Nature Reviews Neuroscience*, 24(9), 557–574. <https://doi.org/10.1038/s41583-023-00718-5>, PubMed: 37438433

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*, 60(2), 378–389. <https://doi.org/10.1016/j.neuron.2008.09.023>, PubMed: 18957228

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of learning and memory*, 82(3), 171–177. <https://doi.org/10.1016/j.nlm.2004.06.005>, PubMed: 15464402

Stent, G. S. (1973). A physiological mechanism for Hebb's postulate of learning. *Proceedings of the National Academy of Sciences*, 70(4), 997–1001. <https://doi.org/10.1073/pnas.70.4.997>, PubMed: 4352227

Sun, W., Advani, M., Spruston, N., Saxe, A., & Fitzgerald, J. E. (2023). Organizing memories for generalization in complementary learning systems. *Nature Neuroscience*, 26(8), 1438–1448. <https://doi.org/10.1038/s41593-023-01382-9>, PubMed: 37474639

Sutherland, R. J., & Lehmann, H. (2011). Alternative conceptions of memory consolidation and the role of the hippocampus at the systems level in rodents. *Current Opinion in Neurobiology*, 21(3), 446–451. <https://doi.org/10.1016/j.conb.2011.04.007>, PubMed: 21592780

Suzuki, A., Josselyn, S. A., Frankland, P. W., Masushige, S., Silva, A. J., & Kida, S. (2004). Memory reconsolidation and extinction have distinct temporal and biochemical signatures. *Journal of Neuroscience*, 24(20), 4787–4795. <https://doi.org/10.1523/JNEUROSCI.5491-03.2004>, PubMed: 15152039

Tonegawa, S., Pignatelli, M., Roy, D. S., & Ryan, T. J. (2015). Memory engram storage and retrieval. *Current Opinion in Neurobiology*, 35, 101–109. <https://doi.org/10.1016/j.conb.2015.07.009>, PubMed: 26280931

Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4(3),

374–391. <https://doi.org/10.1002/hipo.450040319>, PubMed: 7842058

Wammes, J., Norman, K. A., & Turk-Browne, N. (2022). Increasing stimulus similarity drives nonmonotonic representational change in hippocampus. *eLife*, 11, e68344. <https://doi.org/10.7554/eLife.68344>, PubMed: 34989336

Watrous, A. J., Tandon, N., Conner, C. R., Pieters, T., & Ekstrom, A. D. (2013). Frequency-specific network connectivity increases underlie accurate spatiotemporal memory retrieval. *Nature Neuroscience*, 16(3), 349–356. <https://doi.org/10.1038/nn.3315>, PubMed: 23354333

Westphal, A. J., Wang, S., & Rissman, J. (2017). Episodic memory retrieval benefits from a less modular brain network organization. *Journal of Neuroscience*, 37(13), 3523–3531. <https://doi.org/10.1523/JNEUROSCI.2509-16.2017>, PubMed: 28242796

Wiedermann, M., Donges, J. F., Kurths, J., & Donner, R. V. (2017). Mapping and discrimination of networks in the complexity-entropy plane. *Physical Review E*, 96(4), 042304. <https://doi.org/10.1103/PhysRevE.96.042304>, PubMed: 29347608

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society*, 17(5), 766–780. <https://doi.org/10.1017/S1355617711000683>, PubMed: 21729403

Winocur, G., Moscovitch, M., & Bontempi, B. (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia*, 48(8), 2339–2356. <https://doi.org/10.1016/j.neuropsychologia.2010.04.016>, PubMed: 20430044