A Neural State-Space for Episodic Memories

Aidan J. Horner[1,2]

1. Department of Psychology, University of York, UK
2. York Biomedical Research Institute, University of York, UK

Prof Aidan J. Horner,
Department of Psychology,
University of York, Heslington, York,
YO10 5DD, U.K.
Website: www.aidanhorner.org
E-mail: aidan.horner@york.ac.uk

Abstract

Episodic memories are highly dynamic and change in non-linear ways over time. This dynamism isn't captured by existing systems consolidation theories that predict a unidirectional process where memories are first supported by the hippocampus and then the neocortex. I propose a three-dimensional state-space for episodic memories. The first two dimensions relate to whether episodic retrieval is driven by the hippocampus and the neocortex, critically allowing for independent and additive contributions from both regions. The third dimension relates to the episodic-specificity of retrieval. Memories can be located at any point in this state-space and move to any other location. The state-space captures the dynamic nature of episodic memory and broadens the search space of possible memory states and transformations across time.

Keywords

Episodic memory, systems consolidation, hippocampus, neocortex, reinstatement

Acknowledgements

Thanks to Jamie Cockcroft for help with the simulations and plotting in Figure 2, and to Gareth Gaskell, Scott Cairney, Dan Denis, and Harriet Over for helpful feedback on an earlier draft of this paper. AJH is funded by the ESRC (ES/X00791X/1) and Leverhulme Trust (RPG-2023-003).

The dynamic nature of episodic memory

Episodic memory, our ability to recollect previous life events, is highly dynamic and reconstructive [1–4]. Once encoded, memories change, and their expression relies on different neural mechanisms and brain regions. To explain this dynamism, multiple theories of **systems consolidation** (Glossary) have been proposed that capture certain aspects of this process (Box 1; [5–12]), such as the potential time-limited role of the hippocampus in supporting episodic memory. However, these theories fail to fully capture the dynamic and reconstructive nature of episodic memory. For example, the role of the hippocampus may not always be time-limited and may remerge as a critical structure [13,14], reversing the predicted hippocampal-to-neocortical transfer of regions that supports recollection. This dynamism and flexibility has recently been noted [15,16], however we lack a theoretical framework for dealing with such dynamism. I propose a new 'state-space' framework to understand how episodic memories are expressed and how we can track changes to this expression over time. Critically, the proposed neural state-space for episodic memories allows for dynamic changes in memory expression over time that most theories of systems consolidation are not able to explain.

I first outline current theories of episodic memory retrieval and systems consolidation before challenging some key assumptions of theories of systems consolidation, highlighting how most theories are not able to fully capture the dynamic nature of episodic memory over time. I use this critique to introduce a new framework – a neural state-space for episodic memory – that allows for the retrieval of memories to dynamically change over time. I demonstrate how existing theories sit within this state-space and outline experimental approaches that would allow for an exploration of this neural state-space.

Neocortical reinstatement underpins episodic recollection

When we experience an event, the individual elements of the event such as where we are and who we are with are represented in patterns of neural activity in distinct neocortical regions (Figure 1). These independently represented elements are bound into an **memory trace** in the hippocampus [17–20], allowing for the retrieval of the event later in time. If I return to the same location I visited recently, the hippocampus retrieves the other elements of the event via **pattern completion** [21–23], driving the **reinstatement** of that information in the neocortex [24–27].

However, the hippocampus does not directly project to all neocortical regions where reinstatement occurs [28–30]. Instead, reinstatement is driven by a process where neocortical representations are 'reconstructed' via back projections along specific neocortical processing pathways [31,32]. Indeed, reinstatement has been shown in both higher-order [24–27] and primary visual regions [33–36], demonstrating that reinstatement can flow back along the entire processing pathway. This process should occur along multiple processing pathways related to the elements of an event. For example, reinstatement of object and location representations might independently occur via reconstructing representations in the ventral and dorsal visual pathways [37]. Recollection is therefore underpinned by neocortical reinstatement, and this reinstatement occurs via a reconstructive process along neocortical processing pathways.

Current theories of systems consolidation

Theories of systems consolidation propose a time-limited role for the hippocampus in driving this reinstatement [5,7]. When an event is first experienced, no direct connections exist between the neocortically represented event elements, and the hippocampus is needed to bind together and subsequently retrieve these elements. Over time, an active process of systems consolidation is

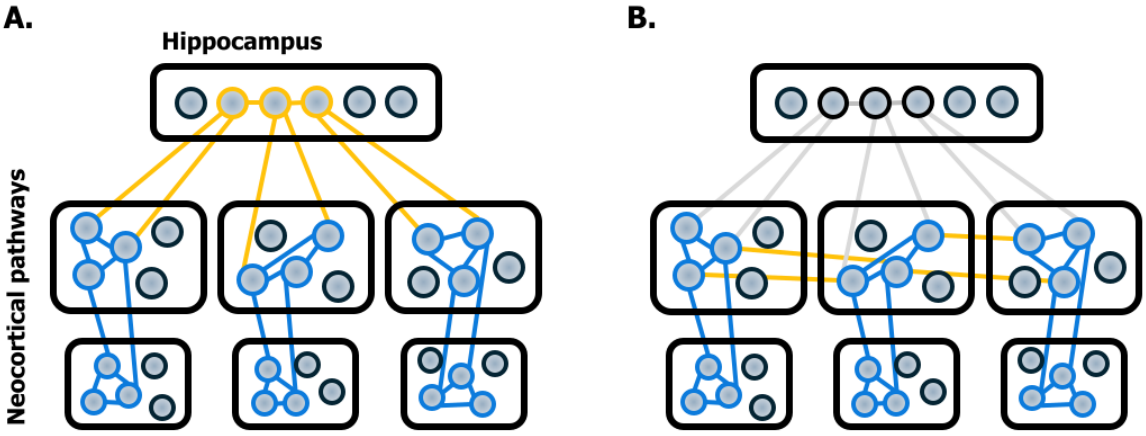

Figure 1. A modal neurocognitive model of episodic retrieval and systems consolidation
A. When experiencing an event, a distributed pattern of activity occurs along neocortical processing pathways (pathways shown here are intentionally 'generic' but could map to ventral and dorsal visual pathways as well as other sensory processing pathways (e.g., auditory)). These pathways convergence on the hippocampus and are bound into a memory trace allowing for their retrieval later in time. At encoding, all neocortical representations are active, forming a new hippocampal memory trace. At retrieval, a subset of neocortical representations will be active, allowing for the retrieval of the hippocampal memory trace and subsequent reinstatement of the remaining neocortical representations. **B.** Over time, connections between neocortical regions form, allowing for their retrieval without the involvement of the hippocampus. Blue lines = pre-existing connections/representations, yellow lines = newly formed connections/representations; grey lines = previously formed connections that have weakened.

thought to occur where connections between the independent neocortical representations form, decreasing the need for the hippocampus to drive reinstatement (Figure 1). These new connections are thought to be formed either directly between neocortical regions [38], or via a second 'hub' such as the ventromedial prefrontal cortex [39]. These connections are predicted to form gradually via the 'reactivation' of hippocampal traces, potentially during offline periods [7] such as non-REM sleep [40,41].

Despite variations across different theories of systems consolidation (Box 1), there is general agreement to the predictions that (1) episodic recollection is underpinned by neocortical reinstatement, (2) neocortical reinstatement is initially driven by the hippocampus, and (3) over time the role of the hippocampus as a driver of reinstatement decreases and another non-hippocampal mechanism drives reinstatement. A final key prediction by specific theories is that over the course of systems consolidation, reinstatement becomes less episodically-specific and more 'gist-like' [8,9]. Critically, recent evidence suggests none of these theories are able to fully capture the dynamic nature of episodic memory.

From hippocampal to neocortical retrieval

Given that recollection is underpinned by the reconstruction of representations along neocortical processing pathways [21,32,37], how do the above tenets of systems consolidation fit with this retrieval mechanism? According to systems consolidation, reinstatement can occur via two mechanisms: (1) the retrieval of a hippocampal memory trace that drives reinstatement (Figure 1A) and/or (2) a non-hippocampal mechanism (Figure 1B). Critically, reinstatement should be less hippocampally driven and more driven by a non-hippocampal mechanism over time.

However, despite this general prediction, theories are rarely explicit in relation to (1) how reinstatement occurs when both mechanisms are available to drive reinstatement and (2) how these two mechanisms are related over time.

How do the two drivers of reinstatement interact at the point of retrieval? Are they compensatory, such that if one fails to drive reinstatement the other takes over, or are they instead additive, such that the strength of reinstatement is driven by the combined contribution of both mechanisms? We have recently provided evidence for an additive relationship – both mechanisms cumulatively contribute to reinstatement. Using fMRI, we correlated neocortical reinstatement with a neural marker of hippocampal pattern completion (hippocampal activity during a memory task where complex events are retrieved in a 'holistic' manner, consistent with the presence of pattern completion; Figure 2). Immediately after encoding, we see a positive relationship – the greater the evidence for hippocampal pattern completion the greater the evidence for neocortical reinstatement [21,42,43]. Importantly, when there is no evidence for hippocampal pattern completion there is no evidence for reinstatement (formally, the intercept of the correlation is zero). This suggests that immediately after encoding the hippocampus is the sole driver of reinstatement.

After 24-hours we still see a significant correlation between the hippocampus and reinstatement, suggesting that the hippocampus is still driving reinstatement, but now also see a non-zero intercept [43]. This positive intercept provides evidence for reinstatement in the absence of

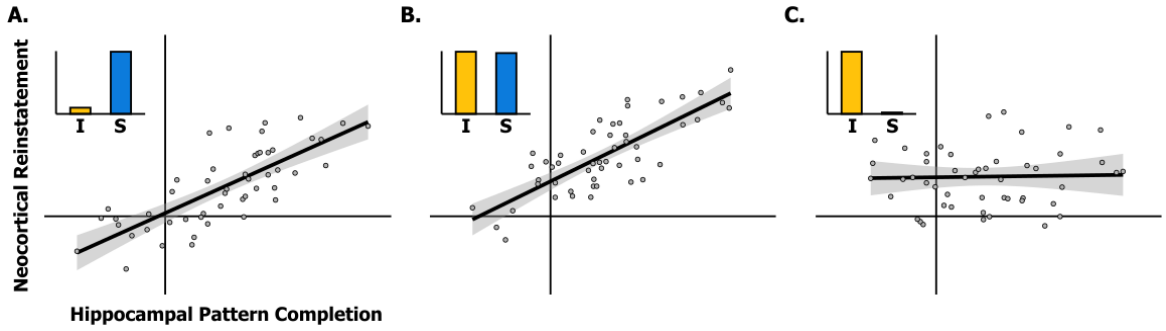

Figure 2. Assessing hippocampal and non-hippocampal contributions to reinstatement
Correlating a marker of hippocampal pattern completion at retrieval (x-axis) with the amount of neocortical reinstatement (y-axis) provides a measure of how much reinstatement is driven by the hippocampus. A positive slope provides evidence for the hippocampus driving reinstatement. A positive intercept provides evidence for the presence of neocortical reinstatement in the absence of evidence for hippocampal pattern completion, suggesting an independent non-hippocampal mechanism is driving reinstatement. Independently assessing the slope and intercept of this relationship therefore provides a measure of the extent to which the hippocampal and non-hippocampal mechanism are driving reinstatement (Slope = hippocampal mechanism; intercept = non-hippocampal mechanism). Each panel shows a pattern of data that could be seen using the experimental approach from [43]. Example data (simulated in R) showing **A.** a positive slope and zero intercept (strong hippocampal, weak non-hippocampal contribution), as seen immediately after encoding in [43], **B.** a positive slope and intercept (strong hippocampal, strong non-hippocampal contribution), as seen 24 hours after encoding in [38], and **C.** zero slope and positive intercept (weak hippocampal, strong non-hippocampal contribution), not seen in [43] but predicted as the end-point of consolidation in most theories of systems consolidation. Critically, the proposed state-space predicts that these patterns could be seen at any timepoint after encoding given specific conditions. See [21,42,43] for real data. I = Intercept, S = Slope in bar graphs.

evidence for hippocampal pattern completion. We infer from this that another, non-hippocampal, mechanism is contributing to reinstatement after 24 hours. Critically, while we see a significant increase in the intercept over 24 hours (from 0 to >0), implying an increase in a non-hippocampal mechanism, we see no evidence for a decrease in the slope of the relationship, implying there is no decrease in the hippocampal mechanism. These data suggest that the two mechanisms additively contribute to neocortical reinstatement – the hippocampus still drives reinstatement to the same degree after 24 hours, but reinstatement is now being boosted by a non-hippocampal mechanism. Further empirical studies need to assess the additive nature of reinstatement, ideally using within-subject correlations (as opposed to the across-subject correlations used in [21,43]) that allow for the assessment of reinstatement at the level of individual episodic memories.

In relation to how the two mechanisms are related over time, Complementary Learning Systems (a computational instantiation of Standard Systems Consolidation Theory) makes the explicit prediction that the non-hippocampal mechanism increases in strength as a function of repeated interleaved reactivation during offline periods [7]. There is ample evidence supporting this prediction [44–48]. The hippocampal mechanism in this model decreases as a consequence of passive decay over time, and recent research supports such a decay mechanism [49], potentially driven via hippocampal **neurogenesis** [50,51]. Thus, the causes of an increase in one mechanism and a decrease in the other are mechanistically distinct – reactivation vs decay.

Current empirical research suggests that the two mechanisms that drive reinstatement are independent and additively contribute to reinstatement and the increase/decrease of these mechanisms are independent. Further, if they are mechanistically distinct, then the time course with which they increase/decrease can potentially differ. They may be loosely temporally correlated, but that does not mean they are mechanistically related to one another nor follow the same precise time-course. The conclusion is that two independent mechanisms with different time-courses additively contribute to neocortical reinstatement.

From detail-rich to gist-like retrieval

A second key concept is the process of semanticisation, where episodic memories become more generalised or 'gist-like' as a function of systems consolidation. This is specifically a prediction of Trace Transformation Theory (TTT) [8,9,12]. Under this theory, the hippocampus supports episodically-specific, detail-rich representations and the neocortex supports more abstract, gist-like representations.

There is evidence showing decreased specificity of the remembered event over time, particularly for the specific perceptual and/or peripheral aspects of the event [52,53]. However, it is not clear whether this more rapid forgetting of event details is mechanistically linked to systems consolidation. If recollection is underpinned by the reconstruction of neocortical representations, then it is not clear why the hippocampus should drive more detail-rich reinstatement than a non-hippocampal mechanism, particularly given the hippocampus does not have direct projections to lower-level sensory regions [28–30].

We can conceptualise 'episodic specificity' in two ways: in relation to (1) the perceptual specificity of reinstated neocortical representations and (2) the forgetting of certain aspects of the event and not others (i.e., the 'coherence' of memory retrieval). In the case of the former, we might reinstate all event elements, but less specifically. For example, we might forget the apple was green and not red. If perceptual specificity at retrieval is a consequence of how far back along

a given processing pathway reactivation occurs, there is no clear reason that the hippocampus should drive reinstatement further back down processing pathways than a non-hippocampal mechanism. Critically, there is currently no empirical data showing that hippocampally-driven reinstatement occurs further back along a given processing pathway relative to non-hippocampally-driven reinstatement. Thus, there is no reason to expect a loss in perceptual specificity as a function of systems consolidation.

If episodic specificity is a function of forgetting certain aspects of the event and not others, such as forgetting where we found the apple but still remembering the apple was green, there is no reason to assume this should occur as a direct consequence of systems consolidation. While offline reactivation might cause certain direct neocortical connections to form at the expense of others, thereby prioritising some aspects of the event, this forgetting of specific elements could occur independently of the active process of systems consolidation (e.g., via forgetting mechanisms such as interference or decay; [54–58]).

In sum, there is no reason that an active process of offline reactivation should cause a change in what is being reactivated. Two independent mechanisms could drive the reinstatement of the same pattern, which would result in the same degree of episodic-specificity at retrieval. Instead, the loss in specificity might be temporally correlated with systems consolidation but caused by forgetting. We should not assume that a loss in episodic-specificity at retrieval is mechanistically related to systems consolidation simply because it is temporally correlated.

The dynamic nature of episodic memory

I have argued that three processes that often temporally co-occur are not mechanistically related, as is assumed in systems consolidation theories. Neocortical reinstatement can be driven, in an additive manner, by two independent mechanisms and the increase/decrease over time of these mechanisms are independent. Further, the specificity with which reinstatement occurs is related to mechanisms that are theoretically independent of what is driving the reinstatement. For example, a strong input from either a hippocampal or non-hippocampal mechanism could both cause a high degree of episodic specificity.

A further critical assumption in most theories of systems consolidation is that the process is unidirectional – from hippocampally-driven reinstatement to non-hippocampally-driven reinstatement over time. Agreeing with TTT, I challenge this assumption [8,12]. First, the hippocampal mechanism can be maintained over longer periods of time by the application of norepinephrine to the amygdala shortly after initial encoding [59]. Second, the non-hippocampal mechanism does not have to slowly emerge over time but can be engaged rapidly either through retrieval-based mechanisms [60–62], or when events are aligned with our schematic understanding of the world [63,64]. Thus, the time-course of both mechanisms can be substantially altered. More critically, the hippocampal mechanism can be increased [65] or even re-engaged [13,14] as a driver of reinstatement at longer delays, suggesting the typical direction of systems consolidation can be entirely reversed. This hippocampal re-engagement is highlighted in studies of **systems reconsolidation**, where a memory trace that is thought to have undergone systems consolidation, and is therefore hippocampus independent, becomes dependent on the hippocampus again following memory reactivation via a reminder cue [13,14]. Thus, the extent to which the hippocampal and non-hippocampal mechanism contribute to reinstatement can differ both after initial encoding and over time in non-linear and non-unidirectional ways.

Finally, the degree to which reinstatement is episodically-specific is also likely highly dynamic. Although specific episodic detail will often be lost over time, some situations likely allow for maintained specificity at the point of retrieval. For example, you might experience an event with a specific object (e.g., your favourite coffee cup). Given you are then likely to encounter the same coffee cup repeatedly, when retrieving the first event you will be able to reinstate a high degree of perceptual specificity in relation to the cup. Conversely, if you saw lots of different coffee cups in the intervening period, the specificity of reinstatement for the cup would likely be much less detailed and more 'gist-like'. As neocortical reinstatement is a reconstructive process, 'episodic-specificity' is more likely driven by a range of factors, including long-term (personal) semantic knowledge [66,67] and experiences following the critical event, rather than what mechanism is driving reinstatement. Indeed, recent research shows that the hippocampus can drive both specific and gist-like retrieval, a process mediated by hippocampal neurogenesis after initial encoding [68,69], underlining how the specificity of reinstatement can be independent of the mechanisms driving reinstatement. Although TTT allows for dynamic changes in episodic-specificity over time, it does not allow for this specificity to be decoupled from the mechanism driving reinstatement [8].

A Neural State-Space for Episodic Memories

I have argued for at least three independent processes that are often mechanistically linked in systems consolidation theories: (1) the creation and loss of a hippocampal trace that drives reinstatement, (2) the creation and loss of a non-hippocampal mechanism that drives reinstatement, and (3) the episodic-specificity of reinstatement at retrieval. I have also argued that these time-courses can differ markedly dependent on encoding conditions and subsequent experiences. To better conceptualise these processes, I propose a three-dimensional neural 'state-space', where any given episodic event can be located at any given point in this space at the point of retrieval and move to any other point within the space under specific conditions (Figure 3).

The first dimension of the state-space is the strength of the hippocampal trace that drives reinstatement (from weak to strong), the second dimension is the strength of the non-hippocampal mechanism that drives reinstatement (from weak to strong), and the third dimension is the degree to which reinstatement is episodically-specific (from low to high specificity). It is possible that more independent dimensions exist. For example, there may be multiple non-hippocampal mechanisms that drive reinstatement (e.g., both direct long-range neocortical connections and new neocortical 'hubs' including the ventromedial prefrontal cortex, medial parietal cortex, and angular gyrus). Further, episodic-specificity may differ in a number of ways, including (but not limited to) the perceptual specificity of reinstatement and the degree to which certain aspects of an event are remembered (the 'coherence' of episodic retrieval [70–73]). Here, I focus on these three dimensions for theoretical clarity and because of their importance in relation to theories of systems consolidation. I refer to the second dimension as a 'non-hippocampal' mechanism, as it is still an open question as to what is driving non-hippocampal reinstatement (or whether multiple non-hippocampal mechanisms exist). A similar multidimensional memory space has been proposed [16], however the specific dimensions were not formally defined and the emphasis of the framework was related to how memories change to be behaviourally relevant as opposed to the focus here on the neurocognitive mechanisms that support episodic retrieval.

This state-space is not a mechanistic account of systems consolidation. However, I have outlined a model of episodic retrieval that focusses on reconstructive reinstatement, which provides a

mechanistic basis for interpreting movement through this state-space. The state-space of episodic memory is a broader theoretical framework that starts with the assumptions that (1) three processes that are mechanistically linked in most systems consolidation theories are independent, and (2) memories can move from any location in this state-space to any other location (Figure 3A). The advantage of this state-space framework is that it allows for the mapping of current theories of systems consolidation (Figure 3B) and highlights the locations and directions of travel in the state-space that these theories do not predict (Figure 3C). For example, TTT predicts that a recently encoded event will be reinstated by the hippocampus in a detail-rich manner and, following systems consolidation, will often be reinstated by a non-hippocampal mechanism in a more gist-like manner (Box 1). Critically, TTT doesn't allow memories to move to specific locations in the state-space such as reinstatement that is detail-rich occurring via a non-hippocampal mechanism.

This mapping of current systems consolidation theories in the state-space highlights regions of the space that have not been considered. For example, strengthening of the hippocampal and non-hippocampal mechanism would result in both mechanisms additively contributing to reinstatement. It is plausible that offline reactivation, or active retrieval, strengthens both

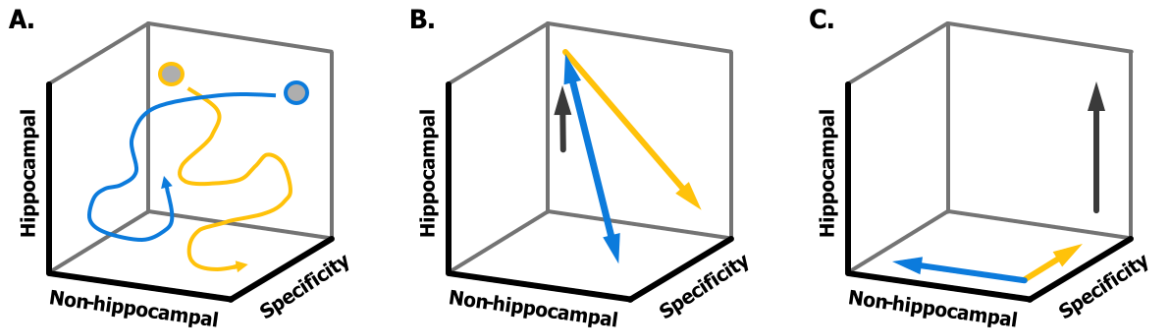

Figure 3. A Neural State-Space for Episodic Memories
The state-space comprises three dimensions – the degree to which reinstatement is driven by a hippocampal mechanism (Hippocampal; weak to strong), the degree to which reinstatement is driven by a non-hippocampal mechanism (Non-hippocampal; weak to strong), and the degree to which reinstatement is episodically-specific (Specificity; low to high). **A. The trajectory of two hypothetical memories.** When an event is experienced it forms a memory trace in the hippocampus (yellow circle; strong hippocampal mechanism) and, depending on the congruence of the event with pre-existing schematic knowledge, can also be supported by the non-hippocampal mechanism (blue circle). Memories take dynamic routes through state-space over time. **B. Theories of systems consolidation mapped as vectors in the state-space.** SSCT starts with strong hippocampal, weak non-hippocampal, and high specificity and moves to weak hippocampal, strong non-hippocampal, and high specificity (yellow arrow). TTT starts with strong hippocampal, weak non-hippocampal, and high specificity and moves to weak hippocampal, strong non-hippocampal, and low specificity (blue arrow). MTT starts with medium hippocampal, weak non-hippocampal, and high specificity and (as more hippocampal traces form) moves to strong hippocampal (the other two dimension remain the same; grey arrow). SCT predicts no movement in this space – the hippocampus is always required for recollection (not shown in subplot B). **C. Underexplored trajectories through the state-space.** Potential trajectories through state-space that are underexplored and/or not currently predicted by existing theories. The forgetting of 'consolidated' memories, from strong non-hippocampal to weak non-hippocampal (blue arrow); an increase in episodic specificity (from low to high) despite reinstatement being driven the non-hippocampal mechanism (yellow arrow); the reintroduction of the hippocampal mechanism (from weak to strong) following 'consolidation' (grey arrow).

mechanisms, leading to memories moving to such a location in the state-space. Indeed, it seems unlikely that the offline reactivation of a hippocampal trace would result in the strengthening of the non-hippocampal mechanism without a similar strengthening of the hippocampal trace itself (Box 2). A further underexplored region of this space is the weakening of the non-hippocampal mechanism. While most theories assume that hippocampal traces decay over time, we do not know how long-range neocortical connections weaken. One possibility is via interference mechanisms that are thought to occur locally in neocortical regions [49,54,57], but we don't know whether interference also occurs for more long-range connections.

By highlighting unexplored regions of this state-space, the idea is to test the strong assumptions of (1) independence and (2) non-unidirectional movement within the state-space. It is possible that these two assumptions don't hold. For example, there might be some areas of the state-space that memories are not able to move to, or certain directions of travel that are not possible. Fully exploring the state-space would allow for the identification of planes or manifolds in this space that memories are constrained to, or attractor states that memories gravitate towards. This broader theoretical framework therefore allows for a wider range of theoretical perspectives, with the intention that fully exploring the state-space will enable the development of more accurate mechanistic models in the future.

Exploring the state-space

One critical question is how best to explore this proposed state-space. How do we test for the assumptions of independence and non-unidirectional movement? First, we need experimental and analysis approaches that can assess the degree to which reinstatement is driven by the hippocampal and non-hippocampal mechanism. Critically, we need to assess how reinstatement is driven at the point of retrieval, which standard univariate analyses that focus on hippocampal and neocortical changes in activity over time are not able to do [74–80]. Our recent approach, assessing the slope and intercept of the relationship between hippocampal pattern completion and neocortical reinstatement is able to do this [43], as are functional connectivity measures between the hippocampus and neocortex [74,81]. Second, we need to assess the degree of 'episodic-specificity' at retrieval. This can be done behaviourally, for example by probing the degree to which participants remember perceptual vs conceptual features of elements in an event [61,82], or more gist-like vs specific features of an event [53]. It can also be achieved using **multivariate pattern analysis** of neuroimaging data, assessing the specificity of reinstatement in neocortical regions [62]. Critically, these approaches would need to be combined, and used at multiple time-points after encoding, to track their proposed independent time-courses.

The above approach would assess the independence of the three state-space dimensions, but not the extent to which movement is non-unidirectional. One direction of travel that SSCT does not currently predict is the re-emergence of the hippocampus in driving reinstatement (Figure 3C). Recent advances in causally manipulating deeper brain regions, for example with **transcranial focussed ultrasound stimulation** [83–85] or **electrical temporal interference stimulation** [86], allow for the up- and down-regulation of the hippocampus. One approach would be to assess the hippocampal contribution to reinstatement at longer delays (where the hippocampus is less likely to be contributing to reinstatement under 'normal' conditions) following upregulation of the hippocampus prior to retrieval. This potential re-emergence of a hippocampal contribution would fit with recent optogenetic research showing hippocampal engrams can be re-activated following a period of forgetting [87]. A further possibility it to use

reminder cues shortly before retrieval following a long delay between encoding and retrieval, promoting systems reconsolidation mechanisms [14].

A second direction of travel that systems consolidation theories do not currently predict is that an older memory that is currently being retrieved with little episodic detail can increase in episodic-specificity (Figure 3C). Here it might be possible to up-regulate specific cortical regions where reinstatement occurs to drive increased specificity, despite the reinstatement being largely driven by a non-hippocampal mechanism. Finally, it might be possible to increase episodic-specificity experimentally with increased exposure to specific event elements across novel events – similar to the example above of using the same coffee cup repeatedly across experiences. Although not exhaustive, I have outlined specific experimental, analysis, and methodological approaches that might allow us to both *track* and *change* the location of an episodic memory in this state-space in humans.

Concluding remarks

Memories dynamically change over time. To account for these changes, theories of systems consolidation predict that the reinstatement of events is first driven by the hippocampus and over time is driven by a non-hippocampal mechanism, and that during this process memories become less episodically-specific. Although all three processes are typically temporally correlated, I have argued that the present evidence suggests that they are mechanistically independent and should be assumed so until evidence to the contrary is provided. I have introduced a 3D state-space to allow, and further test, for this mechanistic independence, where a given episodic memory can be located anywhere in this state-space at any given moment in time. Further, the location of a given memory can move from any location in this state-space to any other. These assumptions of mechanistic independence and non-unidirectional movement challenge current theories of systems consolidation and broaden the search space of possible changes to memories over time. I have further outlined recent experimental, analysis, and methodological advances that allow us to search this proposed state-space, allowing for more mechanistic models of memory change to be developed. More broadly, setting out a wider theoretical space in which current mechanistic models of systems consolidation can be placed allows for better comparison between models and highlights areas of theoretical space that have received less attention or have not been envisaged (see Outstanding Questions). Critically, the state-space framework of episodic memory captures the dynamic nature of memory and provides a theoretical grounding for how this dynamism occurs.

Glossary

Electrical temporal interference stimulation (TIS) – a non-invasive neuromodulation technique that uses multiple electrical fields at different frequencies, allowing for stimulation at superficial and deep (e.g., the hippocampus) brain regions where the fields converge

Memory trace – an interconnected network of neurons in the brain that allows for the retrieval of a previously experienced event at a later time point

Multivariate pattern analysis (MVPA) – the analysis and comparison of patterns of activity in the brain that are associated with particular experimental conditions or mental states

Neurogenesis – the creation of new neurons throughout the lifespan, often seen in the dentate gyrus (DG) of the hippocampus

Pattern completion – the retrieval of a complete neural representation or pattern in the presence of an ambiguous or partial input, commonly associated with subfield CA3 of the hippocampus

Reinstatement – the reestablishment of a similar distributed pattern of neocortical activity during retrieval to that seen when the event was first experienced

Systems consolidation – the reorganisation of a memory trace such that it is supported by different brain regions, typically between an initially acquired hippocampal memory trace to a trace that is less hippocampal dependent.

Systems reconsolidation – the process by which a memory that has undergone systems consolidation to become hippocampus-independent becomes hippocampal-dependent following the reactivation of the memory via a reminder cue.

Transcranial focussed ultrasound stimulation (tFUS) – a non-invasive neuromodulation technique that focusses ultrasound waves to modulate both superficial and deep (e.g., the hippocampus) brain regions

Box 1 – Theories of Systems Consolidation

Standard Systems Consolidation Theory (SSCT) – SSCT proposes that the hippocampus rapidly encodes new episodic experiences and is critical to their retrieval [5,7]. Over time there is a 'transfer' (Box 2) from the hippocampus to neocortex, such that the retrieval of events becomes less hippocampal-dependent over time. This hippocampal-to-neocortical 'transfer' is thought to occur via offline reactivation of hippocampal traces, leading to the gradual formation of non-hippocampal connections between neocortical representations.

Multiple Trace Theory (MTT) – MTT proposes that the hippocampus is always critical to the retrieval of episodic events and that each time an event is retrieved, a new hippocampal trace will be formed creating multiple similar traces of the original experience [10]. The overlapping traces are thought to allow for the extraction of generalities over time, supporting the emergence of semantic knowledge from episodic events.

Trace Transformation Theory (TTT) – TTT builds on SSCT and MTT, proposing that the hippocampus is always critical to the retrieval of episodic events, however overtime the neocortex can support the retrieval of more 'gist-like' or semanticised memories [8,9,12]. Thus, the 'transfer' from the hippocampus to neocortex is accompanied by a change from episodically-specific to more gist-like retrieval. Once both a detail-rich and a gist-like representation has been formed, TTT allows for either trace to dominate at retrieval over time.

Scene Construction Theory (SCT) – SCT proposes that the hippocampus is always needed for the construction of the episodic context or 'scene' of a previous event [11]. While hippocampal traces are predicted to have a relatively short lifespan, a new hippocampal trace is re-established at longer time delays via the reconstruction of consolidated neocortical representations associated with the event. Thus, whereas the elements of an event can be 'consolidated' in the neocortex, the reconstruction of the event into a spatially coherent 'scene' requires the hippocampus regardless of delay.

Box 2 – The fallacy of hippocampal-to-neocortical transfer

The concept of systems consolidation is often associated with a 'transfer' from a hippocampal to a neocortical memory trace. Although perhaps not explicitly stated, there is often an implicit assumption that mnemonic information is transferred from the hippocampus to neocortex. However, the expression of an episodic memory is via the reconstructive reinstatement of representations in the neocortex. As such, there is no 'transfer' of mnemonic information, but instead, at most, a transfer in relation to what mechanism is driving reinstatement. The representational content is always in the neocortex, it is just reactivated via different mechanisms over time.

However, a transfer process implies that one mechanism takes over from the other – as the non-hippocampal mechanism increases the hippocampal mechanism should decrease in an equal manner. Two issues undermine this concept. First, as discussed in the main text, the decrease in hippocampal mechanism is likely driven by an independent process (decay) relative to the process that increases the non-hippocampal mechanism (offline reactivation). Second, it is unclear how the hippocampal mechanism decrease occurs in the presence of continual offline reactivation of the hippocampal trace. Why is reactivation of a hippocampal trace not reinforcing of the trace itself? The default assumption in neuroscience is that synaptic strength increases as a function of neural firing via Hebbian learning. Under this assumption, the reactivation of a hippocampal trace should strengthen itself (as well as increasing any non-hippocampal mechanism). McClelland et al [7] speculate that an additional mechanism might exist, for example the suppression of hippocampal synaptic plasticity during sleep [88] or synaptic downscaling during sleep [89], that supresses this reinforcement. However, such mechanisms would likely only prevent further strengthening and, in the case of synaptic downscaling, would not be trace specific. As such, offline reactivation could lead to the strengthening of both the hippocampal and non-hippocampal mechanism.

Given I have argued for two independent mechanisms that drive reinstatement at retrieval and that the decrease in one mechanism is not related to the increase in the other mechanism, then the notion of 'transfer' is a fallacy. There is no transfer, but simply two distinct drivers of reinstatement that increase/decrease over time via different mechanisms.

References

1. Bartlett, F.C. (1932) *Remembering: A study in experimental and social psychology*, Cambridge University Press
2. Favila, S.E. *et al.* (2020) Transforming the Concept of Memory Reactivation. *Trends Neurosci.* 43, 939–950
3. Schacter, D.L. (2012) Constructive memory: past and future. *Dialogues Clin. Neurosci.* 14, 7–18
4. Staresina, B.P. and Wimber, M. (2019) A Neural Chronometry of Memory Recall. *Trends Cogn. Sci.* DOI: 10.1016/j.tics.2019.09.011
5. Alvarez, P. and Squire, L.R. (1994) Memory consolidation and the medial temporal lobe: A simple network model. *Proc. Natlional Acad. Sci.* 91, 7041–7045
6. Kumaran, D. *et al.* (2016) What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. *Trends Cogn. Sci.* 20, 512–534
7. 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
8. Sekeres, M.J. *et al.* (2018) The hippocampus and related neocortical structures in memory transformation. *Neurosci. Lett.* 680, 39–53
9. Winocur, G. and Moscovitch, M. (2011) Memory Transformation and Systems Consolidation. *J. Int. Neuropsychol. Soc.* 17, 766–780
10. Nadel, L. and Moscovitch, M. (1997) Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr. Opin. Neurobiol.* 7, 217–227
11. Barry, D.N. and Maguire, E.A. (2019) Remote Memory and the Hippocampus: A Constructive Critique. *Trends Cogn. Sci.* 23, 128–142
12. Winocur, G. *et al.* (2010) Memory formation and long-term retention in humans and animals: convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia* 48, 2339–2356
13. Lei, B. *et al.* (2024) Reconstructing a new hippocampal engram for systems reconsolidation and remote memory updating. *Neuron* 0
14. Debiec, J. *et al.* (2002) Cellular and Systems Reconsolidation in the Hippocampus. *Neuron* 36, 527–538
15. Moscovitch, M. and Gilboa, A. (2022) Has the concept of systems consolidation outlived its usefulness? Identification and evaluation of premises underlying systems consolidation. *Fac. Rev.* 11
16. Roüast, N.M. and Schönauer, M. (2023) Continuously changing memories: a framework for proactive and non-linear consolidation. *Trends Neurosci.* 46, 8–19
17. Eichenbaum, H. *et al.* (2007) The Medial Temporal Lobe and Recognition Memory. *Annu. Rev. Neurosci.* 30, 123–152
18. Damasio, A.R. (1989) The Brain Binds Entities and Events by Multiregional Activation from Convergence Zones. *Neural Comput.* 1, 123–132
19. Davachi, L. (2006) Item, context and relational episodic encoding in humans. *Curr. Opin. Neurobiol.* 16, 693–700
20. Marr, D. (1971) Simple Memory: A Theory for Archicortex. *Philos. Trans. R. Soc. Lond. B* 262, 23–81
21. Horner, A.J. *et al.* (2015) Evidence for holistic episodic recollection via hippocampal pattern completion. *Nat. Commun.* 6, 1–11
22. Neunuebel, J.P. and Knierim, J.J. (2014) CA3 Retrieves Coherent Representations from Degraded Input: Direct Evidence for CA3 Pattern Completion and Dentate Gyrus Pattern Separation. *Neuron* 81, 416–427

23. Norman, K. a and O'Reilly, R.C. (2003) Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol. Rev.* 110, 611–646
24. Polyn, S.M. *et al.* (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310, 1963–1966
25. Johnson, J.D. *et al.* (2009) Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. *Neuron* 63, 697–708
26. Gordon, A.M. *et al.* (2014) Cortical Reinstatement Mediates the Relationship Between Content-Specific Encoding Activity and Subsequent Recollection Decisions. *Cereb. Cortex* 24, 3350–3364
27. Staresina, B.P. *et al.* (2012) Episodic Reinstatement in the Medial Temporal Lobe. *J. Neurosci.* 32, 18150–18156
28. Reznik, D. *et al.* (2024) Phylogeny of neocortical-hippocampal projections provides insight in the nature of human memory. *eLife* 13
29. Preston, A. and Wagner, A. (2007) The medial temporal lobe and memory. In *Neurobiology of Learning & Memory*, pp. 305–337, Elsevier
30. Lavenex, P. and Amaral, D.G. (2000) Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus* 10, 420–430
31. Kerrén, C. *et al.* (2025) Exploring the role of dimensionality transformation in episodic memory. *Trends Cogn. Sci.* 29, 614–626
32. Linde-Domingo, J. *et al.* (2019) Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat. Commun.* 10, 179
33. Bosch, S.E. *et al.* (2014) Reinstatement of Associative Memories in Early Visual Cortex Is Signaled by the Hippocampus. *J. Neurosci.* 34, 7493–7500
34. Bone, M.B. and Buchsbaum, B.R. (2021) Detailed Episodic Memory Depends on Concurrent Reactivation of Basic Visual Features within the Posterior Hippocampus and Early Visual Cortex. *Cereb. Cortex Commun.* 2, tgab045
35. Bone, M.B. *et al.* (2020) Feature-specific neural reactivation during episodic memory. *Nat. Commun.* 11, 1945
36. Favila, S.E. *et al.* (2022) Perception and memory have distinct spatial tuning properties in human visual cortex. *Nat. Commun.* 13, 5864
37. Staresina, B.P. *et al.* (2013) Reversible Information Flow across the Medial Temporal Lobe: The Hippocampus Links Cortical Modules during Memory Retrieval. *J. Neurosci.* 33, 14184–14192
38. Klinzing, J.G. *et al.* (2019) Mechanisms of systems memory consolidation during sleep. *Nat. Neurosci.* 22, 1598–1610
39. Frankland, P.W. and Bontempi, B. (2005) The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130
40. Marshall, L. *et al.* (2006) Boosting slow oscillations during sleep potentiates memory. *Nature* 444, 610–613
41. Ngo, H.-V.V. *et al.* (2013) Auditory Closed-Loop Stimulation of the Sleep Slow Oscillation Enhances Memory. *Neuron* 78, 545–553
42. Grande, X. *et al.* (2019) Holistic Recollection via Pattern Completion Involves Hippocampal Subfield CA3. *J. Neurosci. Off. J. Soc. Neurosci.* 39, 8100–8111
43. Joensen, B.H. *et al.* (2024) An Enduring Role for Hippocampal Pattern Completion in Addition to an Emergent Nonhippocampal Contribution to Holistic Episodic Retrieval after a 24 h Delay. *J. Neurosci.* 44
44. Cairney, S.A. *et al.* (2018) Memory Consolidation Is Linked to Spindle-Mediated Information Processing during Sleep. *Curr. Biol.* 28, 948-954.e4
45. Rasch, B. *et al.* (2007) Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science* 315, 1426–1429

46. Rudoy, J.D. *et al.* (2009) Strengthening Individual Memories by Reactivating Them During Sleep. *Science* 326, 1079
47. Schreiner, T. *et al.* (2021) Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nat. Commun.* 12, 3112
48. Staresina, B.P. *et al.* (2013) Awake reactivation predicts memory in humans. *Proc. Natl. Acad. Sci.* 110, 21159–21164
49. Sadeh, T. *et al.* (2016) Forgetting Patterns Differentiate Between Two Forms of Memory Representation. *Psychol. Sci.* 27, 810–820
50. Epp, J.R. *et al.* (2016) Neurogenesis-mediated forgetting minimizes proactive interference. *Nat. Commun.* 7, 10838
51. Frankland, P.W. *et al.* (2013) Hippocampal neurogenesis and forgetting. *Trends Neurosci.* 36, 497–503
52. Fisher, J.S. and Radvansky, G.A. (2018) Patterns of forgetting. *J. Mem. Lang.* 102, 130–141
53. Sekeres, M.J. *et al.* (2016) Recovering and preventing loss of detailed memory: Differential rates of forgetting for detail types in episodic memory. *Learn. Mem.* 23, 72–82
54. Andermane, N. *et al.* (2021) Forgetting across a hierarchy of episodic representations. *Curr. Opin. Neurobiol.* 67, 50–57
55. Hardt, O. *et al.* (2013) Decay happens: The role of active forgetting in memory. *Trends Cogn. Sci.* 17, 111–120
56. Richards, B.A. and Frankland, P.W. (2017) The Persistence and Transience of Memory. *Neuron* 94, 1071–1084
57. Sadeh, T. *et al.* (2014) How we forget may depend on how we remember. *Trends Cogn. Sci.* 18, 26–36
58. Wixted, J.T. (2004) The Psychology and Neuroscience of Forgetting. *Annu. Rev. Psychol.* 55, 235–269
59. Atucha, E. *et al.* (2017) Noradrenergic activation of the basolateral amygdala maintains hippocampus-dependent accuracy of remote memory. *Proc. Natl. Acad. Sci.* 114, 9176–9181
60. Antony, J.W. *et al.* (2017) Retrieval as a Fast Route to Memory Consolidation. *Trends Cogn. Sci.* 21, 573–576
61. Lifanov, J. *et al.* (2021) Feature-specific reaction times reveal a semanticisation of memories over time and with repeated remembering. *Nat. Commun.* 12, 3177
62. Ferreira, C.S. *et al.* (2019) Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information. *NeuroImage* 201, 115996
63. Tse, D. *et al.* (2007) Schemas and Memory Consolidation. *Science* 316, 76–82
64. van Kesteren, M.T.R. *et al.* (2010) Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc. Natl. Acad. Sci.* 107, 7550–7555
65. Krenz, V. *et al.* (2021) Noradrenergic arousal after encoding reverses the course of systems consolidation in humans. *Nat. Commun.* 12, 6054
66. Renoult, L. *et al.* (2019) From Knowing to Remembering: The Semantic – Episodic Distinction. *Trends Cogn. Sci.* DOI: 10.1016/j.tics.2019.09.008
67. Renoult, L. *et al.* (2012) Personal semantics: at the crossroads of semantic and episodic memory. *Trends Cogn. Sci.* 16, 550–558
68. Ko, S.Y. *et al.* (2025) Systems consolidation reorganizes hippocampal engram circuitry. *Nature* DOI: 10.1038/s41586-025-08993-1
69. Golbabaei, A. *et al.* (2025) Neurogenesis-dependent transformation of hippocampal memory traces during systems consolidation. *Curr. Biol.* DOI: 10.1016/j.cub.2025.09.005
70. Horner, A.J. and Burgess, N. (2013) The associative structure of memory for multi-element events. *J. Exp. Psychol. Gen.* 142, 1370–1383

71. Horner, A.J. and Burgess, N. (2014) Pattern Completion in Multielement Event Engrams. *Curr. Biol.* 24, 988–992
72. Hicks, J.L. and Starns, J.J. (2015) Using Multidimensional Encoding and Retrieval Contexts to Enhance Our Understanding of Stochastic Dependence in Source Memory. *Psychol. Learn. Motiv.* 62, 1–40
73. Meiser, T. *et al.* (2008) Binding of multidimensional context information as a distinctive characteristic of remember judgments. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 32–49
74. Takashima, A. *et al.* (2009) Shift from Hippocampal to Neocortical Centered Retrieval Network with Consolidation. *J. Neurosci.* 29, 10087–10093
75. Niki, K. and Luo, J. (2002) An fMRI Study on the Time-Limited Role of the Medial Temporal Lobe in Long-Term Topographical Autobiographic Memory. *J. Cogn. Neurosci.* 14, 500–507
76. Takashima, A. *et al.* (2006) Declarative memory consolidation in humans: A prospective functional magnetic resonance imaging study. *Proc. Natl. Acad. Sci.* 103, 756–761
77. Maguire, E.A. *et al.* (2001) Activity in prefrontal cortex, not hippocampus, varies parametrically with the increasing remoteness of memories. *NeuroReport* 12, 441
78. Ryan, L. *et al.* (2001) Hippocampal complex and retrieval of recent and very remote autobiographical memories: Evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus* 11, 707–714
79. Addis, D.R. *et al.* (2004) Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Hippocampus* 14, 752–762
80. Du, X. *et al.* (2019) Differential activation of the medial temporal lobe during item and associative memory across time. *Neuropsychologia* 135, 107252
81. Cowan, E. *et al.* (2020) Sleep Spindles Promote the Restructuring of Memory Representations in Ventromedial Prefrontal Cortex through Enhanced Hippocampal–Cortical Functional Connectivity. *J. Neurosci.* 40, 1909–1919
82. Andermane, N. *et al.* (2025) The holistic forgetting of events and the (sometimes) fragmented forgetting of objects. *Cognition* 255, 106017
83. Murphy, K. and Fouragnan, E. (2024) The future of transcranial ultrasound as a precision brain interface. *PLOS Biol.* 22, e3002884
84. Yaakub, S.N. *et al.* (2023) Transcranial focused ultrasound-mediated neurochemical and functional connectivity changes in deep cortical regions in humans. *Nat. Commun.* 14, 5318
85. Wessel, M.J. *et al.* (2023) Noninvasive theta-burst stimulation of the human striatum enhances striatal activity and motor skill learning. *Nat. Neurosci.* 26, 2005–2016
86. Violante, I.R. *et al.* (2023) Non-invasive temporal interference electrical stimulation of the human hippocampus. *Nat. Neurosci.* 26, 1994–2004
87. Ryan, T.J. *et al.* (2015) Engram cells retain memory under retrograde amnesia. *Science* 348, 1007–1013
88. Leonard, B.J. *et al.* (1987) Suppression of hippocampal synaptic plasticity during slow-wave sleep. *Brain Res.* 425, 174–177
89. Tononi, G. and Cirelli, C. (2003) Sleep and synaptic homeostasis: a hypothesis. *Brain Res. Bull.* 62, 143–150