The hippocampus as a visual area organized by space and time: A spatiotemporal similarity hypothesis

Nicholas B. Turk-Browne*

Department of Psychology, Yale University, USA

ABSTRACT

The hippocampus is the canonical memory system in the brain and is not typically considered part of the visual system. Yet, it sits atop the ventral visual stream and has been implicated in certain aspects of vision. Here I review the place of the hippocampal memory system in vision science. After a brief primer on the local circuitry, external connectivity, and computational functions of the hippocampus, I explore what can be learned from each field about the other. I first present four areas of vision science (scene perception, imagery, eye movements, attention) that challenge our current understanding of the hippocampus in terms of its role in episodic memory. In the reverse direction, I leverage this understanding to inform vision science in other ways, presenting a working hypothesis about a unique form of visual representation. This spatiotemporal similarity hypothesis states that the hippocampus represents objects according to whether they co-occur in space and/or time, and not whether they look alike, as elsewhere in the visual system. This tuning may reflect hippocampal mechanisms of pattern separation, relational binding, and statistical learning, allowing the hippocampus to generate visual expectations to facilitate search and recognition.

1. Introduction

This article originated in an address I gave at the Vision Sciences Society (VSS) in 2016. I spoke about some surprising ways in which attention and perception interact with long-term memory (Aly & Turk-Browne, 2016a; Hindy, Ng, & Turk-Browne, 2016). Past experience and perceptual learning are known to shape visual processing over a long timescale, but here I was referring to a potential role for the hippocampus in vision — a brain region traditionally thought to be dedicated to episodic, declarative memory.

The status of long-term memory and the hippocampus in vision science remains controversial and elusive. For example, at the next VSS meeting in 2017, only five abstracts out of hundreds reported an interest in, or findings from, the hippocampus. One possible explanation for this disconnect is that the hippocampus is not involved in a meaningful way in visual tasks or phenomena. Another possibility is that the hippocampus is involved but gets ignored for the sociological reason that it is viewed as part of a different brain system (medial temporal lobe, MTL) and as being the concern of a different field (memory research). The gap may also be conceptual, as standard frameworks for understanding the computations and tuning of visual areas do not apply naturally to the hippocampus.

The purpose of this article is to re-introduce the hippocampus to vision scientists and to explore the interface of these fields. This exploration reveals certain aspects of vision that engage the hippocampus, even in cases that do not seemingly involve long-term memory, challenging our current understanding of the hippocampus. In turn, this understanding, rooted in spatial navigation in rodents and episodic memory in humans, suggests a unique potential role for the hippocampus in the visual processing hierarchy.

2. Primer on the hippocampus

General knowledge about the human hippocampus often does not extend far beyond Henry Molaison, the classic case of a patient who suffered amnesia after hippocampal resection (Corkin, 2013). In that spirit, here I provide a modern primer on the circuitry and function of the hippocampus from the human memory literature. This will help clarify why the findings discussed next — of hippocampal involvement in visual tasks — do not readily fit with our current understanding. It will also provide the foundation for later thinking about what can be learned about vision from these known principles of hippocampal function.

* Address: Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06520, USA.
 E-mail address: nicholas.turk-browne@yale.edu.

2.1. Local circuitry

The hippocampus and surrounding MTL cortex constitute one of the best understood systems in the human brain, largely because it is relatively well-conserved across species. As a result, human researchers have been able to draw upon remarkable progress in characterizing the substructures, connectivity, and function of the rodent and non-human primate hippocampus.

MTL cortex is the primary source of input to the hippocampus and also helps translate hippocampal output back into cortical systems. It comprises three primary areas: parahippocampal cortex (PHC; known as postrhinal cortex in rodents), perirhinal cortex (PRC), and entorhinal cortex (ERC). PHC and PRC are most interconnected with the visual system and exhibit partial category selectivity for scenes and objects, respectively (Barense, Henson, Lee, & Graham, 2010; Davachi, 2006; Ranganath & Ritchey, 2012); though this difference is often discussed in terms of PHC representing "contexts" and PRC "items". PHC and PRC both project to ERC (primarily medial and lateral aspects, respectively), and ERC in turn provides input to the hippocampus via superficial layers and receives output from the hippocampus on deep layers. This output is recycled into the superficial layers creating recurrence (Koster et al., 2018) and is sent back down the cortical hierarchy to allow for reinstatement of retrieved content.

Within the hippocampus, there are four main subfields (Fig. 1): dentate gyrus (DG), cornu ammonis area 3 (CA3), CA1, and subiculum (Deng, Aimone, & Gage, 2010; Shohamy & Turk-Browne, 2013; Small, Schobel, Buxton, Witter, & Barnes, 2011). They are connected to form two key pathways: the trisynaptic pathway (TSP; sometimes called perforant pathway), which connects ERC to DG, CA3, CA1, back to ERC (including via the subiculum); and the monosynaptic pathway (MSP; sometimes called temporoammonic pathway), which connects ERC to CA1 and back to ERC; CA3 also has auto-associative connections to itself, allowing for local recurrence and auto-associative binding.

2.2. External connectivity

This circuitry within the hippocampus enables unique and powerful computations, but connectivity with other brain systems is what makes the hippocampus so central to many aspects of behavior. There are strong connections via MTL cortex from (and to) visual areas including inferior temporal (IT) cortex and V4. For instance, learning in IT — which codes object identity invariant to viewpoint, size, etc. — can be impaired by downstream lesions in PRC (Higuchi & Miyashita, 1996), which would also block output from the hippocampus. Beyond canonical visual areas, the MTL is interconnected with frontal and parietal lobes, the amygdala, and via the fornix, the thalamus, mammillary bodies, and basal forebrain (Bird & Burgess, 2008). There are also connections with other brain systems involved in learning and memory, such as the striatum and nucleus accumbens (Pennartz, Ito, Verschure, Battaglia, & Robbins, 2011), neuromodulatory systems, including for dopamine (Shohamy & Adcock, 2010), and sensory systems of all modalities (Lavenex & Amaral, 2000).

The hippocampus is not a homogenous structure, containing both subfield divisions and differences along the anterior-posterior axis (ventral-dorsal in rodents). These latter longitudinal differences manifest in the nature of processing within the hippocampus and in the connectivity with other brain regions (Ranganath & Ritchey, 2012). The anterior hippocampus is more connected to PRC, which in turn serves as a critical hub of an anterior-temporal network of regions including the amygdala, temporal pole, and orbitofrontal cortex. The posterior hippocampus is more connected to PHC, which serves as the hub of a different, posterior-medial network including the retrosplenial cortex, precuneus, angular gyrus, and ventromedial prefrontal cortex.

These distributed systems allow for richer integration with cortex. The anterior-temporal network represents individual entities such as objects or faces (Barense et al., 2010; Buckley & Gaffan, 2006), extracts their value and social relevance (Olson, Plotzker, & Ezzyat, 2007), and links them up with conceptual features (Patterson, Nestor, & Rogers, 2007). The posterior-medial network represents scenes and places (Epstein, Parker, & Feiler, 2007; Park, Intraub, Yi, Widders, & Chun, 2007), tracks temporal order and context (Norman & Eacott, 2005; Turk-Browne, Simon, & Sederberg, 2012), and links these up with situation models such as schemas and event scripts (Baldassano, Hasson, & Norman, 2018).

2.3. Computational principles and models

What does the circuitry and connectivity of the hippocampus do for cognition? The primary function is often considered to be episodic encoding. This refers to our ability to form a detailed memory of a specific moment in space and time (e.g., where I parked my car this morning). There are two interesting challenges for this kind of memory. First, it needs to be "one-shot" in that you only truly experience an episode once; even if a situation is re-encountered in the future with
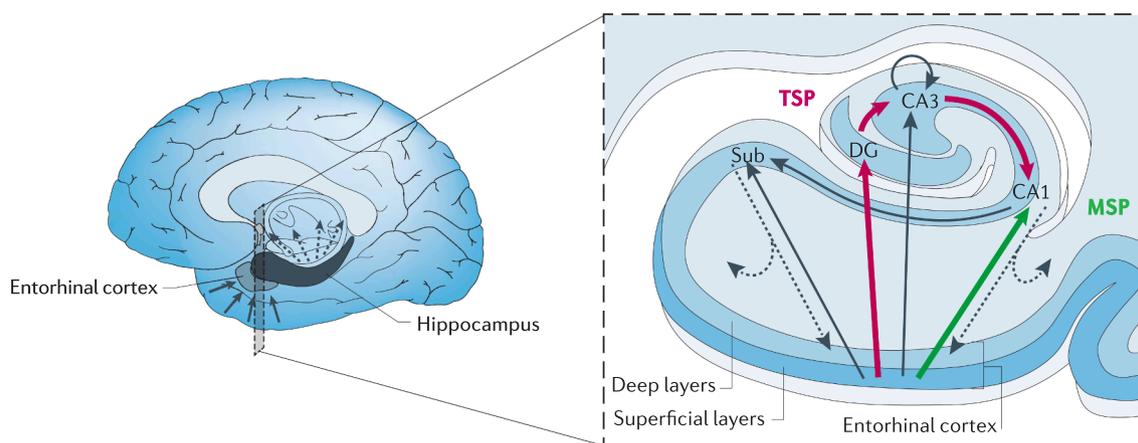
Fig. 1. Location and circuitry of the hippocampus. The hippocampus in each hemisphere is located on the medial aspect of the temporal lobe. It receives input from, and sends output to, several cortical and subcortical systems. The primary local circuitry is depicted here in a coronal section. Inputs (solid arrows) arrive in the hippocampus on superficial layers of entorhinal cortex (ERC) and are routed along both the trisynaptic pathway (TSP; magenta), through dentate gyrus (DG) and CA3 to CA1, and the monosynaptic pathway (MSP; green), directly to CA1. Outputs (dashed arrow) of both pathways arrive on deep layers of ERC, including via the subiculum (Sub), which project to cortical and subcortical targets. Adapted with permission from Small et al. (2011). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

seemingly identical sensory features, the time is different by definition. Second, it is critical to avoid interference with other memories, especially for related episodes (e.g., where I parked my car yesterday).

The hippocampus is thought to solve these challenges with "pattern separation" in the TSP (Leutgeb, Leutgeb, Moser, & Moser, 2007; Yassa & Stark, 2011). DG and then CA3 employ fast learning and sparse coding, which allows them to form orthogonal representations of related experiences that have similar sensory features and therefore overlapping ERC input patterns (e.g., parking in the same neighborhood, at the same time of day, in the same vehicle, etc.). Such episodic memories are complex, unfurling over space and time with input from different modalities. These different components are linked via auto-associative connections in CA3 (Wallenstein, Hasselmo, & Eichenbaum, 1998), a form of relational binding (Cohen & Eichenbaum, 1993). These links in turn allow memory retrieval via "pattern completion": when a cue (i.e., part of a memory) is encountered (e.g., a street intersection or parking garage elevator), the missing components of the original experience are reactivated associatively in CA3, resulting in their output to CA1 and reinstatement in cortex.

A fundamental problem arises out of the interplay of pattern separation and completion: If every slightly different experience is encoded with an orthogonal representation, how can we ever get back to the original memory in CA3 to enable filling-in of missing details? One possibility is that the hippocampus toggles between encoding and retrieval modes, biased for pattern separation and completion, respectively, achieving both functions in the same system in rapid succession (Duncan, Sadanand, & Davachi, 2012; Hasselmo & Stern, 2014).

The account so far would be complete if all we wanted to store in memory was individual moments. However, many aspects of these episodes are idiosyncratic and unlikely to be encountered again (e.g., a stranger walking by, an unusual sports car in the lot). Thus, in order to use memory adaptively in new situations it is also important to abstract over these details and store regularities that are stable across related episodes. This is incompatible with certain characteristics of the hippocampus that make it well suited for episodic memory: unlike one-shot encoding, such learning must occur gradually so that multiple experiences can be integrated over time; against pattern separation, this integration is only possible if related episodes are assigned overlapping representations whose common elements can be strengthened.

This is the core idea behind the complementary learning systems theory (McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003), which posits that the hippocampus also plays a role in establishing memories elsewhere in the brain. Specifically, the hippocampus encodes individual episodes during the day, and during sleep replays these memories to the cortex in an interleaved fashion that leads to gradual learning of their commonalities while avoiding catastrophic interference. This explains a wide range of behavioral and neural data, including from development and patient studies, and has been extended to account for new findings related to associative inference and schema-based learning in the hippocampus (Kumaran, Hassabis, & McClelland, 2016). It remains controversial whether episodic memories become independent of the hippocampus after cortical consolidation, as posited by standard consolidation theory (Squire, 1992). Alternative accounts such as multiple trace theory (Frankland & Bontempi, 2005; Nadel & Moscovitch, 1997) retain a role for the hippocampus after consolidation, arguing that it orchestrates episodic retrieval by maintaining pointers to where different components of a memory are stored in cortex.

3. Perceptual functions of the hippocampus

Until this point, I have discussed the most widely accepted function of the human hippocampus — episodic memory — and how this function is enabled by computations within the hippocampal circuit. In this section, I review four areas of vision science that have implicated the hippocampus but cannot readily be explained by reference to these known principles. In this way, I hope to highlight how vision science can be productive in helping to constrain and advance a general understanding of the hippocampus.

Over the last decade, there has been a wave of studies suggesting a broader role for the hippocampus in cognition (Henke, 2010; Nadel & Hardt, 2011; Shohamy & Turk-Browne, 2013), including some aspects of visual perception (Nadel & Peterson, 2013). This latter conclusion has been especially controversial and the subject of debate (Baxter, 2009; Lee, Yeung, & Barense, 2012; Suzuki, 2009; Suzuki & Baxter, 2009). The points of contention center on whether neuropsychological patients who show perceptual deficits have damage restricted to the hippocampus, whether apparent perceptual deficits are better described as deficits in perceptual learning, and whether perceptual tasks inadvertently place demands on working memory or long-term memory. I do not intend to resolve this debate, but rather to highlight areas of vision science that would not *prima facie* be considered episodic memory tasks and thus might provide useful fodder for rethinking hippocampal functions and models.

3.1. Scene perception

Perhaps the strongest evidence for hippocampal contributions to vision comes from studies of scene perception. In one study (Lee, Bussey et al., 2005), patients with MTL damage were tested on a scene discrimination task in which they were presented with two scenes morphed to different degrees between two endpoint scenes. The task was to determine which of the morphs was most similar to one of the endpoint scenes presented simultaneously as a third stimulus on the screen. The same task was also used to test discrimination of faces, objects, colors, and art. Patients with focal hippocampal lesions were impaired on scene discrimination, especially when the two scenes that needed to be discriminated were closer morphs. However, performance was intact in these focal patients for the other stimulus classes. In contrast, patients with broader MTL damage showed deficits for scenes, faces, and objects.

This general pattern was confirmed in another study with a different perceptual task that required patients to identify which of four presented scenes did not belong with the others (Lee, Buckley et al., 2005). Again, focal hippocampal damage impaired scene but not face oddity judgments, especially in a more challenging condition where all scenes appeared from different viewpoints. These findings from patients have since been validated and extended with neuroimaging studies of the healthy brain (Barense et al., 2010; Hodgetts et al., 2017). What these studies have in common is the need to represent spatial configurations, to abstract and generalize these representations, and to make subtle relational distinctions. Further patient work has shown that the hippocampus is particularly necessary for perceiving and comparing the relational structure of scenes, in contrast to other kinds of scene discrimination that can be solved by noticing changes in discrete features (Aly, Ranganath, & Yonelinas, 2013).

3.2. Imagery

Beyond perceiving scenes, the hippocampus is also involved in imagery for scenes, also known as scene construction. In such tasks (Hassabis, Kumaran, Vann, & Maguire, 2007), participants are cued verbally with a scenario ("Imagine you're lying on a white sandy beach in a beautiful tropical bay") and asked to imagine and see in their "mind's eye" a new, vivid, and multi-modal experience for that scenario. The imagery of hippocampal patients lacks richness and experiential feel, and the content tends to be spatially disjointed and incoherent. This may be related to the deficits noted above in scene perception. Indeed, the anterior hippocampus is involved in both perceiving and constructing scenes (Zeidman, Mullally, & Maguire, 2015). Moreover, there is a relationship between the amount of detail provided by healthy adults when describing a photograph, constructing a scene

in mind, and retrieving an autobiographical memory (Gaesser, Sacchetti, Addis, & Schacter, 2011).

3.3. Eye movements

The factors that determine how eye movements are deployed have long been of central interest to vision science (Hayhoe & Ballard, 2005; Itti & Koch, 2000; Yarbus, 1967). Memory is known to play an important role, including evidence for guidance by contextual and semantic features (Torralba, Oliva, Castelhano, & Henderson, 2006). Some of the effects of memory on eye movements have been attributed to the hippocampus. For example, when viewing a familiar scene that has been manipulated in some way (e.g., removing or moving an object), healthy participants fixate on the manipulated area even when they lack explicit memory for the scene, but hippocampal amnesics do not (Ryan, Althoff, Whitlow, & Cohen, 2000). Moreover, the extent to which eye movements are guided by implicit memories can be predicted by the amount of concurrent hippocampal activity (Hannula & Ranganath, 2009). Eye movements interact with the hippocampus during encoding as well, with more fixations leading to greater hippocampal activity (Liu, Shen, Olsen, & Ryan, 2017). Interestingly, eye movements are sufficient to activate spatial representations in MTL, even in the absence of memory or navigation task demands (Julian, Keinath, Frazzetta, & Epstein, 2018; Killian, Jutras, & Buffalo, 2012; Nau, Schröder, Bellmund, & Doeller, 2018).

3.4. Attention

Episodic memory behavior can be modulated by both selective attention (Uncapher & Rugg, 2009) and divided attention (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). There are two potential explanations. One possibility is that the mnemonic effects of attention are a byproduct of the standard modulation of sensory cortex by attention. Under a biased competition framework (Desimone & Duncan, 1995), such attentional enhancement may increase the likelihood that visual inputs are represented in downstream areas, including the hippocampus, biasing memory encoding to attended inputs. This is consistent with the reported lack of attentional modulation of evoked responses in the hippocampus (Dudukovic, Preston, Archie, Glover, & Wagner, 2011; Yamaguchi, Hale, D'Esposito, & Knight, 2004). An alternative possibility is that attention directly modulates the hippocampus, but the neural signatures of this modulation are different than in sensory cortex. Consistent with this possibility, manipulating which of two modalities is task-relevant to rodents (i.e., whether visuospatial or olfactory cues predict reward) increases the stability of coding in the hippocampus for features of the attended modality (i.e., place fields vs. scent tuning, respectively) (Muzzio et al., 2009).

We recently discovered a similar mechanism for direct attentional modulation of the human hippocampus (Aly & Turk-Browne, 2016b). During a complex visual search task in a virtual art museum, participants searched a rapid stream for a target scene among distractor scenes. The target was defined by either the spatial configuration of the scene (room state) or the style of paintings on the wall (art state). The pattern of functional magnetic resonance imaging (fMRI) activity over voxels in the hippocampus was more reliable across trials of the same state vs. different state, consistent with the idea that attention stabilizes the hippocampus. Moreover, this stability in a combined CA2/CA3/DG region of interest was selectively correlated with behavioral performance in the attention task for the room state.

To test whether this signature of attention in the hippocampus could be responsible for attentional modulation of memory behavior, we examined whether the stability on a given trial impacts which aspects of an experience are encoded into episodic memory (Aly & Turk-Browne, 2016a). We found that the more the pattern of activity in CA2/CA3/DG during incidental encoding of scenes resembled the canonical state for room or art attention, the more likely participants were to later remember the room or painting from that trial, respectively. This link between attention and memory was selective to CA2/CA3/DG.

Attention and memory interact in the other causal direction as well (Hutchinson & Turk-Browne, 2012), with memory retrieval guiding the deployment of attention in familiar environments. For example, greater hippocampal activity in response to a known scene is associated with faster response times in finding a repeated target during visual search (Stokes, Atherton, Patai, & Nobre, 2012). More generally, the hippocampus is involved in (Greene, Gross, Elsinger, & Rao, 2007) and necessary for forms of contextual cueing (Chun & Phelps, 1999) that facilitate visual search.

4. Spatiotemporal similarity hypothesis

It is unclear how the known principles of hippocampal function reviewed earlier can account for some of the visual behaviors above. In this section I explore whether these principles might nevertheless carry some utility for advancing vision science. The hippocampus has mostly been excluded from models of the visual system, which generally terminate in IT prior to the MTL (Kriegeskorte, 2015; Rousselet, Thorpe, & Fabre-Thorpe, 2004; Yamins & DiCarlo, 2016). This is not an anatomical boundary, however, as there are strong bidirectional connections between IT and the hippocampus via MTL cortex. Indeed, the hippocampus has been depicted as the pinnacle of the visual processing hierarchy in primates (Felleman & Essen, 1991).

Rather, I argue that the boundary between the canonical visual system and the hippocampal system is more conceptual artifact, reflecting an assumption about what counts as visual information. IT (and to some extent PHC and PRC) may be the final stage of the visual processing hierarchy in which neural representations are strongly governed by visual appearance. Two objects sharing invariant features that give rise to the same identity or category (e.g., two lamps, cars, or trees) will be represented in similar distributed patterns of neural activity. The core of my proposal — the *spatiotemporal similarity hypothesis* — is that this differs fundamentally from the hippocampus, where representations discount whether two objects have a similar appearance in favor of whether they tend to co-occur in similar locations in space and moments in time. For example, a specific lamp and the particular desk on which it sits in my office might be represented similarly in the hippocampus, despite the fact that this lamp does not look like this desk. Conversely, two lamps or two desks observed in different places and/or at different times might be represented separately, even if they are visually identical.

The added value of this kind of coding scheme to high-level visual processing, distinguishing the hippocampus from IT and lower-level visual cortex, is that similar appearance is neither necessary nor sufficient for objects to be combined. Rather, integration in the hippocampus depends on the proximity of objects in space and time. In fact, there is some evidence that spatiotemporal similarity may lead to relatively more hippocampal integration when objects begin with *dissimilar* appearance (Favila, Chanales, & Kuhl, 2016; Schapiro, Kustner, & Turk-Browne, 2012). Below I review existing evidence for the spatiotemporal similarity hypothesis, describe two potential mechanisms for it based on known principles of hippocampal function, consider how such tuning can be useful, and discuss current gaps in this perspective.

4.1. Existing evidence

There is growing evidence for elements of the spatiotemporal similarity hypothesis. First, if the hippocampus tracks co-occurrence in time, then it should treat nearby moments in time as more similar than distant moments. This is consistent with findings that patterns of activity in the hippocampus carry information about the temporal position of objects in a learned sequence and that objects in adjacent positions are represented more similarly (Deuker, Bellmund, Navarro Schröder, & Doeller, 2016; Ezzyat & Davachi, 2014; Hsieh, Gruber,

Jenkins, & Ranganath, 2014). Second, if the hippocampus prioritizes co-occurrence over appearance, then it should treat similar-looking visual inputs that occur at different times as dissimilar. This is consistent with findings that the hippocampus responds to visually similar "lure" objects as if they were entirely novel (Bakker, Kirwan, Miller, & Stark, 2008) and that highly overlapping routes from different navigational sessions are represented *less* similarly in the hippocampus than when there is no overlap (Chanales, Oza, Favila, & Kuhl, 2017).

Perhaps the most direct evidence for the spatiotemporal similarity hypothesis comes from studies of statistical learning. In one study (Schapiro et al., 2012), participants were shown a continuous sequence of fractal images. Unbeknownst to them, the sequence contained temporal pairs, with the first fractal in a pair always followed by the second. There were no cues to the boundaries between pairs, so the pairs could only be learned from the transition probabilities between fractals in the sequence. Before and after sequence exposure, the fractals were shown in a random order and fMRI was used to measure how the hippocampus represented each fractal individually. Fractals that were paired in the sequence came to be represented more similarly to each other. The assignment of objects to pairs was arbitrary with respect to visual features, so this increased neural similarity was driven by temporal proximity. The same kind of result was obtained in a study of more complex "community structure" (Schapiro, Turk-Browne, Norman, & Botvinick, 2016), whereby the hippocampus came to represent objects that were randomly assigned to the same temporal community more similarly. In a related study, exposure to a sequence of objects generated from a random walk over an arbitrary graph led to fMRI adaptation in the hippocampus that tracked the distance between nodes in the graph (Garvert, Dolan, & Behrens, 2017).

4.2. Potential mechanisms

Tuning for spatiotemporal similarity — that is, selectivity for objects that appear nearby in space and/or time — may be especially strong in the hippocampus because of its unique circuitry introduced earlier (Fig. 1). The two key pathways in the hippocampus — the trisynaptic pathway (TSP) from ERC- > DG- > CA3- > CA1 and the monosynaptic pathway (MSP) from ERC- > CA1 — have the ability to store co-occurring objects together and to keep these representations distinct from those of similar-looking objects that appear in other contexts. Below I describe the mechanics of each pathway based on a recent computational model (Schapiro, Turk-Browne, Botvinick, & Norman, 2017), building upon a long tradition of biologically plausible modeling of the human hippocampus (McClelland et al., 1995; Norman & O'Reilly, 2003). This will highlight how and when the pathways behave differently, clarifying under which circumstances one or the other pathway might drive spatiotemporal similarity.

The key property of the TSP that allows it to store representations based spatiotemporal similarity is pattern separation. The cortical areas that provide input to the hippocampus are organized according to visual features and appearance, such that two similar-looking objects encountered at different points in time may generate overlapping ERC representations. However, if not virtually identical, the TSP assigns these patterns distinct representations (Leutgeb et al., 2007). This occurs because of high inhibition in DG and CA3 that leads few neurons to be active and thus reduces the probability that these neurons will overlap. This explains why visual appearance does not determine TSP representations. The sensitivity of the TSP to spatiotemporal similarity arises from the fact that objects that co-occur simultaneously or in close succession will be active together in ERC, leading to co-activation in CA3 that allows for associative binding across lateral connections. The result of this learning process is that individual objects will come to elicit similar CA3 representations through spreading activation within that subfield to other objects previously encountered at same time.

The main contribution of the MSP to spatiotemporal similarity is statistical learning. Whereas the TSP memorizes every ERC input

pattern and assigns it a distinct representation, the MSP has a slower learning rate that allows it to integrate over inputs. This can be critical because any given episode will contain a mix of objects whose presence and co-occurrences are stable over time (e.g., the lamp and the desk in my office) and objects that are idiosyncratic and less likely to appear again (e.g., a visiting colleague or a ladder being used for repairs). The stable co-occurrences (or *regularities*) are most important to learn from the perspective of using spatiotemporal similarity to make accurate predictions in the future. However, because of the presence of idiosyncratic objects, the input from ERC will differ slightly across related experiences. This is what prevents TSP from extracting even slightly noisy regularities. In contrast, lower inhibition and thus reduced sparsity in MSP lead to greater overlap in the set of active CA1 neurons for these experiences, which in turn allows the subset of neurons consistently activated (corresponding to the stable aspects of the input) to get reinforced. By gradually learning from related experiences, the MSP extracts regularities to form a conjunctive representation of reliably co-occurring objects in CA1. Individual objects can thus elicit similar CA1 representations by reactivating the same conjunction.

These two potential mechanisms are not mutually exclusive and in fact could operate in parallel. Yet, they differ in some important ways. The first difference is in terms of the timescale of learning: the TSP can learn spatiotemporal similarity extremely quickly, even after a single experience, at the risk of encoding spurious co-occurrences; the MSP learns more slowly, after multiple experiences, with the benefit of extracting reliable co-occurrences. The second difference is in terms of how they handle the appearance-based similarity of objects that do not co-occur in space and time (e.g., my lamp at the office and my lamp at home): strong pattern separation in the TSP ensures that these objects will be stored separately; weaker pattern separation in the MSP, essential for extracting regularities across experiences that contain idiosyncratic details, increases the potential for these objects to activate overlapping CA1 representations. This risk is tempered by the fact that most other details of the contexts in which these objects appear will differ (e.g., office vs. home), resulting in largely distinct ERC representations and reducing the likelihood of integration in MSP.

The ideas above are largely theoretical speculations, though in some cases they are grounded in computational simulations (Schapiro et al., 2017). My hope is that this framework will prompt future investigations into the visual functions of the TSP and MSP, and in particular, the roles of DG/CA3 vs. CA1 in vision tasks. These subregions have become accessible to high-resolution fMRI (Carr, Rissman, & Wagner, 2010; Yushkevich et al., 2015) and increasingly to intracranial recordings in epilepsy patients (Solomon et al., 2019), both of which can be used in a variety of vision tasks.

4.3. Consequences of spatiotemporal tuning

What are the benefits of coding for spatiotemporal similarity in the hippocampus? Such representations can be used to generate expectations: seeing a familiar object may activate an integrated representation in the hippocampus containing other objects with which that object co-occurs, which in turn can lead to reinstatement of those expected objects in visual cortex. This can be viewed as a form of predictive coding (Rao & Ballard, 1999; Spratling, 2010). Just as mid-level visual areas that recognize a shape send feedback to lower-level areas to fill in illusory contours (Kok and de Lange, 2014), the hippocampus as a higher-level area may "recognize" a spatiotemporal cluster of objects from partial input and send feedback to mid- and lower-level areas (MTL cortex, IT, lateral occipital cortex, etc.) to fill-in the missing objects. Indeed, the hippocampus seems to generate expectations about associated objects (Hindy et al., 2016; Kok & Turk-Browne, 2018). These signals are related to, and sometimes precede, expectation signals in visual cortex.

Beyond expectation, coding for spatiotemporal similarity may lead to behavioral predictions about perceived similarity. If two objects

come to be represented as one in the hippocampus when paired in time or space, this may bias judgments about the similarity of their appearance and may even make them more confusable (Schapiro et al., 2012). One potential mechanism for this could be if statistical learning in the hippocampus alters tuning for objects in IT. Such pair coding has been observed previously (Li & DiCarlo, 2008; Miyashita, 1988).

4.4. Limitations of the hypothesis

One potential limitation is that it is unclear whether spatiotemporal similarity is unique to the hippocampus. An alternative theory known as the representational-hierarchical view (Saksida & Bussey, 2010) likewise emphasizes the MTL as an extension of the ventral visual stream, but focuses on the role of PRC rather than hippocampus in forming complex conjunctions of features. This provides a parsimonious account of deficits in object discrimination in amnesic patients and lesioned animals. A possible extension compatible with the spatiotemporal similarity hypothesis is to posit that the hippocampus forms complex conjunctions of *objects* rather than features, especially in scenes (Lee et al., 2012). This could account for spatial discrimination deficits in hippocampal patients. However, it is hard to reconcile this emphasis on space with the statistical learning studies reviewed above that contained only temporal regularities.

Increased representational similarity for objects that co-occur has been observed outside of the hippocampus not only in MTL cortex (Schapiro et al., 2012), but also in IT (Messinger, Squire, Zola, & Albright, 2001; Miyashita, 1988). Moreover, highly similar visual inputs that do not co-occur in space or time can be represented differently not only in the hippocampus, but also in V1 (Saleem, Diamanti, Fournier, Harris, & Carandini, 2018). An interpretation of these findings consistent with the spatiotemporal similarity hypothesis is that spatiotemporal signals originate in the hippocampus and propagate from there through PRC, into IT and even earlier visual areas. This fits with findings that, after paired associate learning, memory signals in PRC precede those in IT (Naya, Yoshida, & Miyashita, 2001), and that PRC is necessary for this kind of learning (Buckley & Gaffan, 2006; Higuchi & Miyashita, 1996). What remains unknown is whether this involvement of the PRC reflects its role as an interface between the hippocampus and IT. Another factor is that the nature of the statistics to be learned seems to determine whether representational changes are selective to the hippocampus or broader in MTL and IT cortices. Co-occurrence that can be extracted based on simple transition probabilities or joint frequency drives broader changes (Schapiro et al., 2012), whereas more complex regularities based on predictive overlap lead to more selective effects in the hippocampus (Schapiro et al., 2016).

Another potential limitation is that representational change in the hippocampus is not always spatiotemporal. The ability to encode the similarity structure along multiple task dimensions is broadly consistent with the notion of a "cognitive map" (Behrens et al., 2018; Schiller et al., 2015; Tolman, 1948). In this framework, spatiotemporal similarity can be thought of as constructing a map based on transitions between features (Constantinescu, O'Reilly, & Behrens, 2016) or objects (Garvert, Dolan, & Behrens, 2017; Schapiro, Turk-Browne, Norman, & Botvinick, 2016) in a metric space. However, the hippocampus can also represent maps for abstract dimensions of a task that are not navigated continuously in space or time. For example, the hippocampus can track the position of agents in a metric space defined by the key social dimensions of "power" and "affiliation" (Tavares et al., 2015). One key difference is that such dimensions and spaces may pre-exist social learning about individual agents, whereas spatiotemporal similarity may be especially important for constructing spaces to begin with.

A related question is whether spatiotemporal similarity applies only to the visual modality or reflects a more general property of the hippocampus that applies to input from any sensory modality. The latter would not be surprising, given that the mechanisms proposed above rely upon hippocampal computations that support episodic memory, itself highly multi-modal. Indeed, rodent hippocampus can represent a map of auditory tones sampled from continuous frequency space (Aronov, Nevers, & Tank, 2017). Nevertheless, given the dominance of the visual modality in human perception and the strong interconnectedness of the hippocampus and visual system, the impact of spatiotemporal similarity in humans may be most profound in vision.

5. Conclusions

The many open questions raised here require much additional work at the intersection of the traditionally separate fields of vision science and memory research. Aspects of vision that engage the hippocampus, such as scene perception, imagery, eye movements, and attention, are currently left unaddressed by hippocampal models, which provides an opportunity to expand and constrain our understanding of this vital brain system. These models in turn provide a well-grounded computational framework for evaluating the roles of different architectures, algorithms, and learning rules in various aspects of high-level vision. The goal of such interdisciplinary efforts would be a more integrated understanding of how the brain supports complex human behaviors that depend upon both perceptual and mnemonic information.

References

Aly, M., Ranganath, C., & Yonelinas, A. P. (2013). Detecting changes in scenes: The hippocampus is critical for strength-based perception. *Neuron, 78*(6), 1127–1137. https://doi.org/10.1016/j.neuron.2013.04.018.

Aly, M., & Turk-Browne, N. B. (2016a). Attention promotes episodic encoding by stabilizing hippocampal representations. *Proceedings of the National Academy of Sciences, 113*(4), E420–E429. https://doi.org/10.1073/pnas.1518931113.

Aly, M., & Turk-Browne, N. B. (2016b). Attention stabilizes representations in the human hippocampus. *Cerebral Cortex, 26*(2), 783–796. https://doi.org/10.1093/cercor/bhv041.

Aronov, D., Nevers, R., & Tank, D. W. (2017). Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature, 543*(7647), 719–722. https://doi.org/10.1038/nature21692.

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science, 319*(5870), 1640–1642. https://doi.org/10.1126/science.1152882.

Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of real-world event schemas during narrative perception. *Journal of Neuroscience, 38*(45), 9689–9699. https://doi.org/10.1523/JNEUROSCI.0251-18.2018.

Barense, M. D., Henson, R. N. A., Lee, A. C. H., & Graham, K. S. (2010). Medial temporal lobe activity during complex discrimination of faces, objects, and scenes: Effects of viewpoint. *Hippocampus, 20*(3), 389–401. https://doi.org/10.1002/hipo.20641.

Baxter, M. G. (2009). Involvement of medial temporal lobe structures in memory and perception. *Neuron, 61*(5), 667–677. https://doi.org/10.1016/j.neuron.2009.02.007.

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron, 100*(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002.

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. *Nature Reviews Neuroscience, 9*(3), 182–194. https://doi.org/10.1038/nrn2335.

Buckley, M. J., & Gaffan, D. (2006). Perirhinal cortical contributions to object perception. *Trends in Cognitive Sciences, 10*, 100–107. https://doi.org/10.1016/j.tics.2006.01.008.

Carr, V. A., Rissman, J., & Wagner, A. D. (2010). Imaging the human medial temporal lobe with high-resolution fMRI. *Neuron, 65*(3), 298–308. https://doi.org/10.1016/j.neuron.2009.12.022.

Chanales, A. J. H., Oza, A., Favila, S. E., & Kuhl, B. A. (2017). Overlap among spatial memories triggers repulsion of hippocampal representations. *Current Biology, 27*(15), 2307–2317.e5. https://doi.org/10.1016/j.cub.2017.06.057.

Chun, M. M., & Phelps, E. A. (1999). Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nature Neuroscience, 2*(9), 844–847. https://doi.org/10.1038/12222.

Cohen, N., & Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System.* Cambridge: MIT Press.

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science, 352*(6292), 1464–1468. https://doi.org/10.1126/science.aaf0941.

Corkin, S. (2013). *Permanent Present Tense: The Unforgettable Life of the Amnesic Patient, H. M.* New York, NY: Basic Books.

Craik, F. I., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology. General, 125*(2), 159–180. https://doi.org/10.1037//0096-3445.125.2.159.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology, 16*(6), 693–700. https://doi.org/10.1016/j.conb.2006.10.012.

Deng, W., Aimone, J. B., & Gage, F. H. (2010). New neurons and new memories: How does adult hippocampal neurogenesis affect learning and memory? *Nature Reviews Neuroscience, 11*(5), 339–350. https://doi.org/10.1038/nrn2822.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205.

Deuker, L., Bellmund, J. L., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *ELife, 5*, e16534. https://doi.org/10.7554/eLife.16534.001.

Dudukovic, N. M., Preston, A. R., Archie, J. J., Glover, G. H., & Wagner, A. D. (2011). High-resolution fMRI reveals match enhancement and attentional modulation in the human medial temporal lobe. *Journal of Cognitive Neuroscience, 23*(3), 670–682. https://doi.org/10.1162/jocn.2010.21509.

Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory's penumbra: Episodic memory decisions induce lingering mnemonic biases. *Science, 337*(6093), 485–487. https://doi.org/10.1126/science.1221936.

Epstein, R. A., Parker, W. E., & Feiler, A. M. (2007). Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *Journal of Neuroscience, 27*(23), 6141–6149. https://doi.org/10.1523/JNEUROSCI.0799-07.2007.

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron, 81*(5), 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042.

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

Felleman, D. J., & Essen, D. C. V. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*(1), 1–47. https://doi.org/10.1093/cercor/1.1.1.

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.

Gaesser, B., Sacchetti, D. C., Addis, D. R., & Schacter, D. L. (2011). Characterizing age-related changes in remembering the past and imagining the future. *Psychology and Aging, 26*(1), 80–84. https://doi.org/10.1037/a0021054.

Garvert, M. M., Dolan, R. J., & Behrens, T. E. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *ELife, 6*, e17086. https://doi.org/10.7554/eLife.17086.001.

Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2007). Hippocampal differentiation without recognition: An fMRI analysis of the contextual cueing task. *Learning & Memory, 14*(8), 548–553. https://doi.org/10.1101/lm.609807.

Hannula, D. E., & Ranganath, C. (2009). The eyes have it: Hippocampal activity predicts expression of memory in eye movements. *Neuron, 63*(5), 592–599. https://doi.org/10.1016/j.neuron.2009.08.025.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences, 104*(5), 1726–1731. https://doi.org/10.1073/pnas.0610561104.

Hasselmo, M. E., & Stern, C. E. (2014). Theta rhythm and the encoding and retrieval of space and time. *NeuroImage, 85*, 656–666. https://doi.org/10.1016/j.neuroimage.2013.06.022.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience, 11*(7), 523–532. https://doi.org/10.1038/nrn2850.

Higuchi, S., & Miyashita, Y. (1996). Formation of mnemonic neuronal responses to visual paired associates in inferotemporal cortex is impaired by perirhinal and entorhinal lesions. *Proceedings of the National Academy of Sciences, 93*(2), 739–743. https://doi.org/10.1073/pnas.93.2.739.

Hindy, N. C., Ng, F. Y., & Turk-Browne, N. B. (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature Neuroscience, 19*(5), 665–667. https://doi.org/10.1038/nn.4284.

Hodgetts, C. J., Voets, N. L., Thomas, A. G., Clare, S., Lawrence, A. D., & Graham, K. S. (2017). Ultra-high-field fMRI reveals a role for the subiculum in scene perceptual discrimination. *Journal of Neuroscience, 37*(12), 3150–3159. https://doi.org/10.1523/JNEUROSCI.3225-16.2017.

Hsieh, L.-T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron, 81*(5), 1165–1178. https://doi.org/10.1016/j.neuron.2014.01.015.

Hutchinson, J. B., & Turk-Browne, N. B. (2012). Memory-guided attention: Control from multiple memory systems. *Trends in Cognitive Sciences, 16*(12), 576–579. https://doi.org/10.1016/j.tics.2012.10.003.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7.

Julian, J. B., Keinath, A. T., Frazzetta, G., & Epstein, R. A. (2018). Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature Neuroscience, 21*(2), 191–194. https://doi.org/10.1038/s41593-017-0049-1.

Killian, N. J., Jutras, M. J., & Buffalo, E. A. (2012). A map of visual space in the primate entorhinal cortex. *Nature, 491*(7426), 761–764. https://doi.org/10.1038/nature11587.

Kok, P., & de Lange, F. P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology, 24*(13), 1531–1535. https://doi.org/10.1016/j.cub.2014.05.042.

Kok, P., & Turk-Browne, N. B. (2018). Associative prediction of visual shape in the hippocampus. *Journal of Neuroscience, 38*(31), 6888–6899. https://doi.org/10.1523/JNEUROSCI.0163-18.2018.

Koster, R., Chadwick, M. J., Chen, Y., Berron, D., Banino, A., Düzel, E., ... Kumaran, D. (2018). Big-loop recurrence within the hippocampal system supports integration of information across episodes. *Neuron, 99*(6), 1342–1354.e6. https://doi.org/10.1016/j.neuron.2018.08.009.

Kriegeskorte, N. (2015). Deep neural networks: A new framework for modeling biological vision and brain information processing. *Annual Review of Vision Science, 1*(1), 417–446. https://doi.org/10.1146/annurev-vision-082114-035447.

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences, 20*(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004.

Lavenex, P., & Amaral, D. G. (2000). Hippocampal-neocortical interaction: A hierarchy of associativity. *Hippocampus, 10*(4), 420–430. https://doi.org/10.1002/1098-1063(2000)10:4<420::AID-HIPO8>3.0.CO;2-5.

Lee, A. C. H., Buckley, M. J., Pegman, S. J., Spiers, H., Scahill, V. L., Gaffan, D., ... Graham, K. S. (2005). Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus, 15*(6), 782–797. https://doi.org/10.1002/hipo.20101.

Lee, A. C. H., Bussey, T. J., Murray, E. A., Saksida, L. M., Epstein, R. A., Kapur, N., ... Graham, K. S. (2005). Perceptual deficits in amnesia: Challenging the medial temporal lobe 'mnemonic' view. *Neuropsychologia, 43*(1), 1–11. https://doi.org/10.1016/j.neuropsychologia.2004.07.017.

Lee, A. C. H., Yeung, L.-K., & Barense, M. D. (2012). The hippocampus and visual perception. *Frontiers in Human Neuroscience, 6*, 91. https://doi.org/10.3389/fnhum.2012.00091.

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science, 315*(5814), 961–966. https://doi.org/10.1126/science.1135801.

Li, N., & DiCarlo, J. J. (2008). Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science, 321*(5895), 1502–1507. https://doi.org/10.1126/science.1160028.

Liu, Z.-X., Shen, K., Olsen, R. K., & Ryan, J. D. (2017). Visual sampling predicts hippocampal activity. *Journal of Neuroscience, 37*(3), 599–609. https://doi.org/10.1523/JNEUROSCI.2610-16.2016.

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.

Messinger, A., Squire, L. R., Zola, S. M., & Albright, T. D. (2001). Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proceedings of the National Academy of Sciences, 98*(21), 12239–12244. https://doi.org/10.1073/pnas.211431098.

Miyashita, Y. (1988). Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature, 335*(6193), 817. https://doi.org/10.1038/335817a0.

Muzzio, I. A., Levita, L., Kulkarni, J., Monaco, J., Kentros, C., Stead, M., ... Kandel, E. R. (2009). Attention enhances the retrieval and stability of visuospatial and olfactory representations in the dorsal hippocampus. *PLoS Biology, 7*(6), e1000140. https://doi.org/10.1371/journal.pbio.1000140.

Nadel, L., & Hardt, O. (2011). Update on memory systems and processes. *Neuropsychopharmacology, 36*(1), 251–273. https://doi.org/10.1038/npp.2010.169.

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.

Nadel, L., & Peterson, M. A. (2013). The hippocampus: Part of an interactive posterior representational system spanning perceptual and memorial systems. *Journal of Experimental Psychology: General, 142*(4), 1242–1254. https://doi.org/10.1037/a0033690.

Nau, M., Schröder, T. N., Bellmund, J. L. S., & Doeller, C. F. (2018). Hexadirectional coding of visual space in human entorhinal cortex. *Nature Neuroscience, 21*(2), 188–190. https://doi.org/10.1038/s41593-017-0050-8.

Naya, Y., Yoshida, M., & Miyashita, Y. (2001). Backward spreading of memory-retrieval signal in the primate temporal cortex. *Science, 291*(5504), 661–664. https://doi.org/10.1126/science.291.5504.661.

Norman, G., & Eacott, M. J. (2005). Dissociable effects of lesions to the perirhinal cortex and the postrhinal cortex on memory for context and objects in rats. *Behavioral Neuroscience, 119*(2), 557–566. https://doi.org/10.1037/0735-7044.119.2.557.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review, 110*(4), 611–646. https://doi.org/10.1037/0033-295X.110.4.611.

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: A review of findings on social and emotional processing. *Brain, 130*(7), 1718–1731. https://doi.

org/10.1093/brain/awm052.

Park, S., Intraub, H., Yi, D.-J., Widders, D., & Chun, M. M. (2007). Beyond the edges of a view: Boundary extension in human scene-selective visual cortex. *Neuron, 54*(2), 335–342. https://doi.org/10.1016/j.neuron.2007.04.006.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976–987. https://doi.org/10.1038/nrn2277.

Pennartz, C. M. A., Ito, R., Verschure, P. F. M. J., Battaglia, F. P., & Robbins, T. W. (2011). The hippocampal–striatal axis in learning, prediction and goal-directed behavior. *Trends in Neurosciences, 34*(10), 548–559. https://doi.org/10.1016/j.tins.2011.08.001.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience, 13*(10), 713–726. https://doi.org/10.1038/nrn3338.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79–87. https://doi.org/10.1038/4580.

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). How parallel is visual processing in the ventral pathway? *Trends in Cognitive Sciences, 8*(8), 363–370. https://doi.org/10.1016/j.tics.2004.06.003.

Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science, 11*(6), 454–461. https://doi.org/10.1111/1467-9280.00288.

Saksida, L. M., & Bussey, T. J. (2010). The representational–hierarchical view of amnesia: Translation from animal to human. *Neuropsychologia, 48*(8), 2370–2384. https://doi.org/10.1016/j.neuropsychologia.2010.02.026.

Saleem, A. B., Diamanti, E. M., Fournier, J., Harris, K. D., & Carandini, M. (2018). Coherent encoding of subjective spatial position in visual cortex and hippocampus. *Nature, 562*(7725), 124–127. https://doi.org/10.1038/s41586-018-0516-1.

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology, 22*(17), 1622–1627. https://doi.org/10.1016/j.cub.2012.06.056.

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Phil. Trans. R. Soc. B, 372*(1711), 20160049. https://doi.org/10.1098/rstb.2016.0049.

Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus, 26*(1), 3–8. https://doi.org/10.1002/hipo.22523.

Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., & Ranganath, C. (2015). Memory and space: Towards an understanding of the cognitive map. *Journal of Neuroscience, 35*(41), 13904–13911. https://doi.org/10.1523/JNEUROSCI.2618-15.2015.

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences, 14*(10), 464–472. https://doi.org/10.1016/j.tics.2010.08.002.

Shohamy, D., & Turk-Browne, N. B. (2013). Mechanisms for widespread hippocampal involvement in cognition. *Journal of Experimental Psychology, 142*(4), 1159–1170. https://doi.org/10.1037/a0034461.

Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P., & Barnes, C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Reviews Neuroscience, 12*(10), 585–601. https://doi.org/10.1038/nrn3085.

Solomon, E. A., Stein, J. M., Das, S., Gorniak, R., Sperling, M. R., Worrell, G., ... Kahana, M. J. (2019). Dynamic theta networks in the human medial temporal lobe support episodic memory. *Current Biology, 29*(7), 1100–1111.e4. https://doi.org/10.1016/j.cub.2019.02.020.

Spratling, M. W. (2010). Predictive coding as a model of response properties in cortical area V1. *Journal of Neuroscience, 30*(9), 3531–3543. https://doi.org/10.1523/JNEUROSCI.4911-09.2010.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99*(2), 195–231. https://doi.org/10.1037/0033-295x.99.2.195.

Stokes, M. G., Atherton, K., Patai, E. Z., & Nobre, A. C. (2012). Long-term memory prepares neural activity for perception. *Proceedings of the National Academy of Sciences, 109*(6), E360–E367. https://doi.org/10.1073/pnas.1108555108.

Suzuki, W. A. (2009). Perception and the medial temporal lobe: Evaluating the current evidence. *Neuron, 61*(5), 657–666. https://doi.org/10.1016/j.neuron.2009.02.008.

Suzuki, W. A., & Baxter, M. G. (2009). Memory, perception, and the medial temporal lobe: A synthesis of opinions. *Neuron, 61*(5), 678–679. https://doi.org/10.1016/j.neuron.2009.02.009.

Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., & Schiller, D. (2015). A map for social navigation in the human brain. *Neuron, 87*(1), 231–243. https://doi.org/10.1016/j.neuron.2015.06.011.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*(4), 189–208. https://doi.org/10.1037/h0061626.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766.

Turk-Browne, N. B., Simon, M. G., & Sederberg, P. B. (2012). Scene representations in parahippocampal cortex depend on temporal context. *Journal of Neuroscience, 32*(21), 7202–7207. https://doi.org/10.1523/JNEUROSCI.0942-12.2012.

Uncapher, M. R., & Rugg, M. D. (2009). Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. *Journal of Neuroscience, 29*(25), 8270–8279. https://doi.org/10.1523/JNEUROSCI.1043-09.2009.

Wallenstein, G. V., Hasselmo, M. E., & Eichenbaum, H. (1998). The hippocampus as an associator of discontiguous events. *Trends in Neurosciences, 21*(8), 317–323. https://doi.org/10.1016/S0166-2236(97)01220-4.

Yamaguchi, S., Hale, L. A., D'Esposito, M., & Knight, R. T. (2004). Rapid prefrontal-hippocampal habituation to novel events. *Journal of Neuroscience, 24*(23), 5356–5363. https://doi.org/10.1523/JNEUROSCI.4587-03.2004.

Yamins, D. L. K., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience, 19*(3), 356–365. https://doi.org/10.1038/nn.4244.

Yarbus, A. L. (1967). *Eye Movements and Vision*. Plenum Press.

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences, 34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006.

Yushkevich, P. A., Amaral, R. S. C., Augustinack, J. C., Bender, A. R., Bernstein, J. D., Boccardi, M., ... Zeineh, M. M. (2015). Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: Towards a harmonized segmentation protocol. *NeuroImage, 111*, 526–541. https://doi.org/10.1016/j.neuroimage.2015.01.004.

Zeidman, P., Mullally, S. L., & Maguire, E. A. (2015). Constructing, perceiving, and maintaining scenes: Hippocampal activity and connectivity. *Cerebral Cortex, 25*(10), 3836–3855. https://doi.org/10.1093/cercor/bhu266.