

Studying Episodic Memory Using Real-time fMRI

October 6, 2021

Jeffrey D. Wammes
Queen's University
Kingston, ON, CA K7M1E7
jeffrey.wammes@queensu.ca

Qi Lin
Yale University
New Haven, CT 06511
qi.lin@yale.edu

Kenneth A. Norman
Princeton University
Princeton, NJ 08544
knorman@princeton.edu

Nicholas B. Turk-Browne
Yale University
New Haven, CT 06511
nicholas.turk-browne@yale.edu

Abstract

Memory refers to our ability to acquire, store, retrieve, and restructure experiences and knowledge over our lifetime. In this way, memory is fundamental to virtually all behaviors and cognitive functions. Although real-time functional magnetic resonance imaging (rtfMRI) has been used to investigate many of these functions, such as perception, attention and emotion, rtfMRI has been used much less to directly study the neural mechanisms underlying memory. In this chapter, we review the existing literature on how the human brain supports memory. We focus on systems that underlie the encoding, consolidation and retrieval of individual experiences in long-term memory, also known as episodic memory. In addition to correlational approaches, we consider interventional studies that have experimentally manipulated the brain to establish causal relationships to memory behavior. Finally, we discuss how rtfMRI could be used to monitor, perturb, and instantiate brain states relevant to memory, covering both recent rtfMRI studies in this nascent field and potential ideas for future research. We conclude that rtfMRI is poised to provide unique theoretical insights into human memory.

1 Introduction

Memory is absolutely critical to our functioning in the world — it allows us to acquire new knowledge, to make generalizations, to develop expertise, to adapt to changes in our world, and to allow prior experience to shape our future behavior. Here, we focus on episodic memory: the ability to form lasting memories of individual experiences (i.e., episodes) that support cued recall of details from those experiences — for example, being able to recall the particular dish that you ordered when you visited a restaurant on a particular day. Understanding the complex brain states that support episodic memory has been a popular research objective. Many of these studies use fMRI as a non-invasive technique that is able to record from deep-brain structures important for episodic memory (e.g., the hippocampus) and that provides the spatial resolution needed to understand how memories are represented in the brain. In most fMRI studies of episodic memory, participants learn some new information in the laboratory and their memory for this information is later tested. This takes place while brain images are being collected every few seconds. Following the experiment, the brain images are compiled, preprocessed and analyzed. The conclusions of such studies are often correlational: better memory in condition A is associated with the presence of brain state P, or conditions B and C can be distinguished based on memory-related information carried in brain region Q. Such analyses provide important insights about how the human brain supports memory, but without directly speaking to the causal role of brain states or regions.

In contrast, rtfMRI can be considered an intervention, affording researchers the unprecedented ability to monitor, perturb and even instantiate brain states noninvasively in healthy, normally functioning humans. Using this approach, the experiment can be turned on its head: in a conventional fMRI design, the tasks or stimuli are predetermined and brain activity is treated as a dependent variable; in an rtfMRI design, brain activity controls the tasks or stimuli and thus serves as more of an independent variable. This can be accomplished in two ways: In a *triggering* design, brain activity determines the timing of a task or stimulus. For example, a stimulus appears when the brain enters into a desired state. In a *neurofeedback* design, brain activity determines the content of the task or stimulus. For example, the activity of a brain region might be conveyed as explicit feedback on a visual gauge. In both cases, the impact of rtfMRI is usually measured in terms of changes in subsequent behavior and brain activity. Conclusions drawn using rtfMRI are not entirely causal, because controlling activity in a brain region and observing a change in subsequent behavior does not rule out additional links in the causal chain (e.g., an intermediate region or an explicit strategy). Nevertheless, the relationships established with rtfMRI are closer to establishing causation than conventional fMRI approaches because an intervention is applied to perturb brain states, as opposed to simply observing how the brain responds to a task manipulation.

Causal approaches have long been possible in non-human animals, but are generally more invasive than is feasible or ethical for humans (e.g., lesion, pharmacology, optogenetics). Some non-invasive approaches used in humans, including electroencephalography (EEG) neurofeedback or transcranial magnetic stimulation (TMS), have supported causal inferences (e.g., Antony et al., 2018; Tambini et al., 2018; Wang et al., 2019). However, these approaches have limitations when studying the mechanisms underlying memory, related to the fact that memories are distributed within and across brain regions (Johnson et al., 2009; Kuhl et al., 2011; Kuhl et al., 2012; Polyn et al., 2005; Ward et al., 2013; Xue et al., 2010). Measuring and manipulating distributed representations requires a technique with: (1) high spatial resolution, (2) the ability to record simultaneously from multiple regions, and (3) access to deep-brain structures central to memory (e.g., hippocampus). EEG and TMS do not satisfy these prerequisites. EEG has high temporal resolution and measures simultaneously across multiple channels, but has low spatial resolution and is most sensitive to brain structures on the cortical surface. TMS can achieve higher spatial specificity (Hanlon, 2017), but is still an order of magnitude coarser than fMRI. TMS is also typically only applied to one or a small number of sites, and there is no direct access to deep-brain structures (though indirect access may be possible: Hermiller et al., 2020; Tambini et al., 2018; Wang et al., 2014).

Invasive methods for studying memory in humans, such as intracranial EEG (iEEG) and direct electrical stimulation (DES), can be highly precise, allow for direct measurement and manipulation of multiple sites, and provide access to deep-brain structures. However, an obvious limitation to this approach is that such procedures are only possible when clinically indicated, for example in patients with intractable epilepsy. Not only is the pool of such patients limited, but electrodes are generally implanted in brain regions because of the potential role of these regions in the disease. This may restrict the generalizability of the results to the healthy brain. Because the placement of electrodes is entirely determined by clinical need, there is also no experimenter control in which brain regions are measured or manipulated. Many electrodes are typically implanted, but the resulting coverage is still far from spanning the whole brain.

fMRI represents a desirable alternative to EEG, TMS, and iEEG/DES for studying the distributed brain states subserving memory. It provides simultaneous measurement of the whole brain, including deep-brain structures, with relatively high spatial resolution. When analyzed in real-time, fMRI can be used for online monitoring and perturbation of the brain in healthy, normally functioning humans. rtfMRI studies build on a strong foundation from conventional fMRI and other techniques, which have identified clear signatures of memory processes that can be targeted for enhancement.

A key part of this foundation is the realization that memory is intrinsically dynamic — the product of several processes interacting on different timescales. When we have an experience, we immediately *encode* it into a memory that persists after the experience ends, then *consolidate* this memory into a more durable state alongside other existing memories, and all the while we can *retrieve* and update the memory based on subsequent experience (Figure 1A). In this chapter, we will review evidence for how these three stages of memory are supported by distributed networks in the human brain. Building on this rich literature, we will consider interventional methods that complement more correlational approaches (e.g. conventional fMRI and EEG) and open up new avenues for developing treatments to enhance or suppress memories. Informed by these existing approaches, in a final section we will discuss the unique opportunities that rtfMRI presents for

causal experimental investigations of episodic memory, through monitoring, perturbing and instantiating brain states.

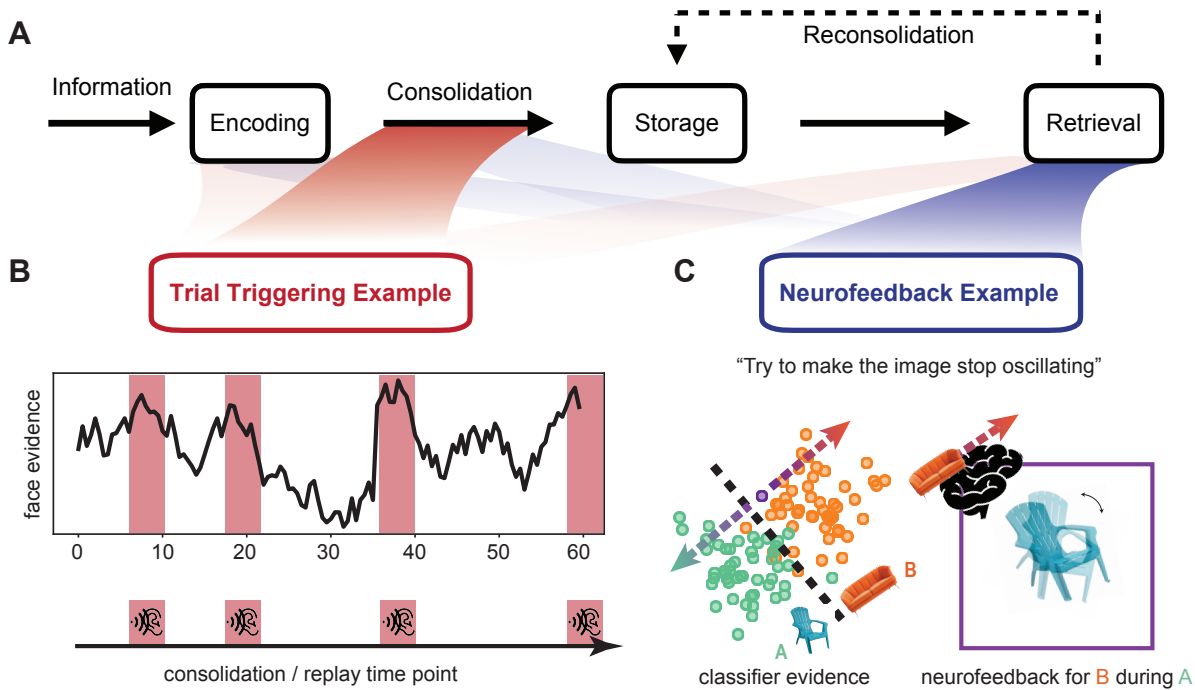


Figure 1: Stages of memory and types of rtfMRI. (A) As we experience the world, this information is internalized through *encoding*. Over time, encoded information is stabilized through *consolidation* to allow for more permanent storage in memory. These stored memories can later come to mind spontaneously or through cues from the environment, resulting in *retrieval*. Every time we retrieve information, it goes through a process of destabilization and restabilization called reconsolidation. Each of these stages can be investigated with rtfMRI, generally in one of two designs: triggering and neurofeedback. (B) For example, triggering rtfMRI designs can be used to study consolidation. During periods of rest or sleep, classifier evidence can be used to detect *replay* of encoded stimuli. When these states are detected, a sound can be played to disrupt replay and thus impair memory (Schreiner et al., 2015). Varying the timing of the disruption and the brain region used for classification could clarify the nature of this stabilization process. (C) As another example, neurofeedback rtfMRI designs can be used to study retrieval. Consider a pair of objects (e.g., chair and couch). When retrieving one member of the pair cued on the screen (i.e., chair), the visual display can be driven by classifier evidence for the other member (i.e., couch). In the depicted example, the chair is wobbling, but this slows as classifier evidence for the couch increases. As a result, the representations of chair and couch are concurrently active, which may lead to their integration in memory (Ritvo et al., 2019; Schlichting & Preston, 2015).

2 Encoding

We are confronted with a barrage of sensory input as we navigate the world. This input is filtered through our perceptual pathways and then must then be recorded in some way so as to inform later behavior in a similar situation or context. Encoding refers to the process whereby these perceived events are converted into an enduring memory trace that can be recalled at a later time (Tulving, 1985). This process is not infallible — many factors can degrade the quality of encoding, including the state of the brain at the time of exposure to new information (Daselaar et al., 2009). Understanding the dynamic nature of the brain states

that underlie both effective and ineffective encoding could drive the development of neurofeedback-based interventions designed to induce desirable brain states, establishing a causal link between these states and encoding.

2.1 Neural Basis and Measures of Encoding

One of the most prominent tasks for studying the dynamic neural states that drive memory encoding is the *subsequent memory paradigm* (Brewer et al., 1998; Wagner et al., 1998). During an initial encoding phase, participants are presented with a series of stimuli (e.g. words or pictures) while their brains are scanned using fMRI. This is followed by a retrieval test, in which participants are asked to judge whether stimuli are “old” or “new”. Stimuli from the encoding phase correctly identified as old are labeled as remembered and those incorrectly reported as new are labeled as forgotten. These labels are used to sort the fMRI trials from the encoding phase. Brain activity can be compared between trials with stimuli that were later remembered and trials with stimuli that were later forgotten. This paradigm has yielded a plethora of important discoveries about brain states that are predictive of successful encoding (as indicated by later retrieval).

Many of these results came from univariate analyses in which activity is collapsed across an entire region of interest (ROI) in the brain. This kind of analysis can determine whether an ROI is more active in a particular task or condition. In subsequent memory studies, frontoparietal cortex, MTL cortex, and the hippocampus consistently show greater activity for remembered relative to forgotten encoding trials, regardless of the type of materials encoded (Kim, 2011; Uncapher & Wagner, 2009). In fact, even before the presentation of a stimulus to be encoded, brain states in MTL cortex and the hippocampus reliably predict subsequent memory (Guderian et al., 2009; Park & Rugg, 2010; Turk-Browne et al., 2006). These encoding states can be considered to be stimulus-agnostic, as the stimulus is not yet known to the brain. Other regions are more sensitive to specific types of input. For example, activity in left inferior frontal cortex predicts memory for words, while activity in bilateral fusiform cortices predicts memory for pictures (Kim, 2011; Uncapher & Wagner, 2009). Together, these findings highlight that the most effective encoding state is likely to recruit content-general regions in concert with distributed regions that are more finely tuned to the specific nature of the incoming content.

Memory research has also benefited from the use of multivariate pattern analysis (MVPA; Lewis-Peacock and Norman, 2014). In contrast to the univariate approach of averaging activation across voxels in a region, MVPA exploits the spatial pattern of activity within a region to determine the content or information being represented there. MVPA has been applied in tasks where to-be-remembered information is presented repeatedly during an encoding phase. Researchers measured the neural patterns at each repetition of the same stimulus, then quantified how stable these patterns were across presentations. During encoding, higher pattern similarity across repetitions of the same event was positively associated with performance in later memory tests (Ward et al., 2013; Xue et al., 2010). This effect is most prominent in frontoparietal cortex, visual cortex, and posterior cingulate cortex.

Univariate and multivariate fMRI analyses reveal the role of individual regions in encoding, but how do these regions interact to support this process? Functional connectivity analyses, which measure the correlation of fMRI activity between pairs of brain regions, provide a window into these network dynamics. Successful encoding is generally associated with increased connectivity between frontoparietal cortex, MTL cortex, and the hippocampus (Ranganath et al., 2005; Summerfield et al., 2006). Connectivity among these regions predicts variance in subsequent memory within and across individuals. These analyses also allow the identification of distinct brain networks supporting different aspects or types of memory. For example, resting-state fMRI (i.e., no cognitive task) has been used to identify brain networks with distinct patterns of MTL connectivity (Kahn et al., 2008; Libby et al., 2012): the posterior medial network, composed of regions interconnected with posterior hippocampus (including parahippocampal cortex, retrosplenial cortex, posterior cingulate cortex, and precuneus), and the anterior temporal network, composed of regions interconnected with anterior hippocampus (including perirhinal cortex, anterior temporal cortex, amygdala, and orbitofrontal cortex). These networks have distinct roles, with the anterior temporal network processing items and the posterior medial network processing context (see Ranganath and Ritchey, 2012).

Successful encoding therefore relies on the the dynamic interaction of spatial patterns of neural activity across distinct networks of brain regions. The most prominent regions identified for encoding processes are the hippocampus, MTL cortex, and frontal cortex, but regions more specific to the type of information

encoded are also engaged (e.g., premotor cortex, fusiform gyrus, and posterior parietal cortex). By measuring whole-brain fMRI during encoding and linking these neural measures to later behavioral memory judgments, it is possible to identify which brain states are conducive to memory. Such findings lay the foundation for interventional studies designed to perturb these states to promote or sabotage encoding.

2.2 Methodological Approaches for Perturbing Encoding

Interventions have been used to influence activity in the brain regions and networks underlying memory encoding, often with the intention of nudging brain states either toward or away from optimal encoding states during a memory task. For example, a TMS study (Tambini et al., 2018) indirectly targeted the hippocampus by stimulating posterior inferior parietal cortex (pIPC), a region robustly connected with the hippocampus, during the encoding of object-location associations. This led to a selective enhancement in associative memory (see also Wang et al., 2014), which is primarily supported by the hippocampus, but not item memory, which can be supported by cortical regions outside of the hippocampus. Although TMS cannot directly target deeper brain structures, this study elegantly circumvented this constraint, establishing that pIPC stimulation likely impacted hippocampal activity. Resonating with that work, a later concurrent fMRI-TMS study convincingly demonstrated that delivering TMS to the parietal cortex during encoding led to an immediate impact on hippocampal activity and improved memory formation (Hermiller et al., 2020). Other TMS studies targeted left dorsolateral prefrontal cortex (PFC) during encoding of verbal and non-verbal material and observed downstream effects on memory (Gagnon et al., 2010; Gagnon et al., 2011).

Unlike TMS, DES can directly perturb the hippocampus and other deep-brain structures. Due to its invasive nature, however, DES is only conducted in patients when it is medically necessary. In epilepsy, electrodes are implanted in the brain for 1-2 weeks to monitor seizures and identify their source location prior to neurosurgery. During this time, electrical potentials are recorded and can be linked to experimental tasks; this method is often referred to as intracranial EEG (iEEG). Near the end of monitoring period, current is applied to the same electrodes to stimulate the neural tissue where they terminate. Because electrodes are often implanted in the MTL cortex and hippocampus (common sites of seizures and sclerosis in epilepsy), these patients provide a rare opportunity to probe for causal effects in memory.

DES applied in the MTL cortex and hippocampal formation leads to variable effects (see Suthana and Fried, 2014). This highlights the complexity of these regions and the distributed nature of the representations they encode. For instance, stimulating entorhinal cortex (a subregion of MTL cortex) during encoding has been shown to enhance spatial memory, while stimulating the hippocampus proper had no measurable impact (Suthana et al., 2012). A later study reported that stimulation of either entorhinal cortex or hippocampus during encoding impaired both verbal free recall and spatial memory performance (Jacobs et al., 2016). In the spatial memory task, participants were placed randomly in an open arena with distant landmarks and automatically moved toward a target that wasn't visible until they started moving. That is, there was no consistent relationship between the person's initial position and the invisible target, preventing an egocentric map and biasing participants toward an allocentric perspective (i.e. based on spatial relationships between targets and external boundaries). Stimulation only impaired spatial memory for locations farther away from boundaries and memory for the temporal ordering of words, both of which require keeping track of relations, the specialty of the hippocampus (Goyal et al., 2018).

Expanding on these results and mirroring the fMRI studies discussed earlier, a recent study showed that subsequent memory can be predicted from hippocampal iEEG recordings acquired during encoding (Ezzyat et al., 2017). This provided a target for perturbing encoding, which was pursued with DES in a later session with the same patients. Namely, during the encoding of a word list, DES was applied to the hippocampus and MTL cortex when possible. This stimulation caused neither an overall increase nor decrease in memory performance. Rather, an interaction was observed, helping to clarify the earlier inconsistent results: When the brain was already in a good encoding state (based on iEEG signal decoding), stimulation disrupted this state and impaired memory; however, when the brain was in a bad encoding state, stimulation nudged the brain toward a better encoding state, which in turn improved free recall performance (Ezzyat et al., 2017). This confirms the existence and causal role of encoding states as an explanation for variability in memory within a person over time. There is now also evidence for variability in the impact of stimulation on memory across individuals (Kucewicz et al., 2018).

Temporal dynamics in encoding states can be leveraged for closed-loop memory improvement using DES

(Ezzyat et al., 2018). In this work, pattern classifiers were trained on iEEG recording data from a patient to distinguish encoding states likely to yield later retrieval or forgetting. The patient then underwent an additional memory task with stimulation applied during encoding. However, rather than randomly assigning trials or time periods for stimulation, the pre-trained classifiers were used to parse incoming data and detect bad encoding states (i.e., those associated with forgetting). When detected, these states were perturbed using stimulation to the lateral temporal cortex. Stimulation disrupted the negative encoding states, yielding improved memory not only for items present during stimulation, but also for items presented nearby in time to the stimulation.

Together, these results indicate that manipulations of a wide range of regions can affect encoding, and that the impact of stimulation depends on the currently unfolding brain state of the individual. In other words, to be confident about when and how to perturb brain states, it is necessary to monitor brain states in real-time. Applying rtfMRI to influence memory performance is a new frontier, but one early study demonstrated the promise of this approach (Yoo et al., 2012). Participants were tasked with encoding a set of scenes. In an initial (non-real-time) study, the researchers established that low pre-stimulus activation in the parahippocampal place area (PPA) was associated with improved subsequent memory. In the next study, PPA activity during encoding was measured in real time. When low PPA activation was detected, presentation of a new scene was triggered. In other words, researchers identified a state that was conducive to encoding in real time and supplied new information for the participant to learn during that time. As a control, other trials were triggered when high PPA activation was detected. Indeed, scenes triggered by low but not high PPA activity were later better remembered. Although this study did not use neurofeedback to change the participant’s brain state, it identified optimal or sub-optimal encoding states, and determined the timing of encoding opportunities accordingly.

3 Consolidation

Once new information has been encoded into memory, it must be accommodated alongside existing memories and maintained in a durable state for later retrieval. Consolidation is the time-dependent process whereby newly encoded information is stabilized and stored more permanently in long-term memory. This process can be discussed at two different levels of description, the cellular/synaptic level and the systems level (Dudai, 2012), which unfold along different timescales. Synaptic consolidation occurs in the local neural circuit that is engaged in encoding the memory, and refers to the time-dependent stabilization of changes in synaptic efficacy in this circuit that often occur within seconds after encoding. In contrast, systems consolidation occurs in a more distributed set of regions across the brain, and refers to the re-organization of neural circuits to allow the more permanent storage of the new information. Here, we focus exclusively on the latter process of systems consolidation, which can be studied with fMRI.

3.1 Neural Basis and Measures of Consolidation

Theories about systems consolidation primarily focus on the interaction between the hippocampus and neocortex. According to *standard consolidation theory* (Marr et al., 1991), newly formed memories are initially dependent on a network of brain regions comparable to those implicated during encoding, including the hippocampus, MTL cortex, and neocortical areas involved in the experience. The hippocampus serves as a fast learner, rapidly binding the features of an experience. In contrast, the neocortex is a slow learner, building longer-lasting but less detailed representations while avoiding interference with existing memories. Specifically, following encoding, the hippocampal state associated with the memory is repeatedly reinstated during sleep or rest, and this repetition through replay leads to gradual changes in neocortical circuits. Once completed, the memories no longer depend upon the hippocampus. Standard consolidation theory was initially supported by findings from amnesic patients who had hippocampal lesions but mostly intact neocortex (Corkin, 2013). This pattern of pathology is associated with a preserved ability to recall semantic knowledge and older memories that have undergone consolidation, but sweeping impairments in the acquisition of new episodic memories.

Later studies found evidence that hippocampal lesions can also lead to deficits for older episodic memories, indicating that remote memory does not always become independent from the hippocampus (see Frankland and Bontempi, 2005). This can be accounted for by *multiple trace theory* (Nadel & Moscovitch, 1997), in which hippocampal replay leads to multiple memory traces for the same information. That is, the hippocampus

stores a “pointer” or “index” to a specific distributed pattern of cortical activity. Because similar traces reference highly overlapping neocortical networks, their accumulation produces abstract, gist-like long-term representations. However, these memories remain anchored in the hippocampus such that the loss of this index after hippocampal lesion renders the neocortical traces inaccessible. The *complementary learning systems* (CLS) theory (McClelland et al., 1995) can also account for these results. According to CLS, hippocampal replay trains cortex to represent abstractions across multiple replayed episodes, rather than simply transferring individual episodic memories to cortex; as such, memory for individual episodes will continue to depend on hippocampus, unless they are replayed very extensively.

The traditional view of consolidation posits that it occurs once, after which the consolidated memory is considered fixed (McGaugh, 1966). However, if a consolidated memory is cued for retrieval, it can become labile again and subject to change, prior to a period of *reconsolidation* (Misanin et al., 1968; Nader et al., 2000). Although many studies of reconsolidation focus on fear conditioning (Nader et al., 2000; Schiller et al., 2010), reconsolidation has since been reported in several memory protocols (see Dudai, 2012; Nader and Hardt, 2009). These and other studies of consolidation highlight the dynamic nature of memory traces: they are formed, replayed, stored, reactivated, and modified repeatedly over time. With this in mind, targeted manipulations that reinstate brain states productive for memory may benefit consolidation and strengthen memories.

3.2 Methodological Approaches for Perturbing Consolidation

Although manipulating the consolidation process directly is less common than manipulating encoding and retrieval, there have been a few attempts using DES. For example, researchers have stimulated memory-relevant brain structures during the consolidation period of a verbal free-recall task (Merkow et al., 2017). DES to MTL cortex during the retention period between encoding and test of a word list impaired later retrieval. Although DES also impaired memory when applied during encoding and retrieval itself, the disruption was greatest during consolidation. This was interpreted as DES interfering with the successful formation of an enduring memory trace, perhaps by preventing replay or rehearsal. Stimulating the amygdala immediately following learning of neutral images can also impact later memory, in a positive direction (Inman et al., 2018). Namely, amygdala stimulation during consolidation increased coherence between the hippocampus and perirhinal cortex during retrieval the following day and improved recognition memory performance.

EEG can also be used to monitor for states that are conducive to memory consolidation. In studies using *targeted memory reactivation* (TMR), sleeping participants are reminded of specific episodes of learning from when they were awake, via presentation of sensory cues associated with those episodes (see Paller and Norman, 2020). These reminders impact hippocampal activity and subsequent memory more during slow wave sleep (SWS) than during rapid eye movement (REM) sleep or wake (e.g., Rasch et al., 2007). SWS can be distinguished from REM using EEG; thus, by monitoring for SWS and presenting cues when this sleep stage is detected, one can optimize the impact on consolidation.

TMR has also been used to show that sleep reminders are more effective at promoting consolidation when they elicit sleep spindle oscillations (Antony et al., 2018; Schreiner et al., 2015; Wang et al., 2019). This association between sleep spindles and memory consolidation presents an excellent target for intervention, made possible by the ability to detect sleep spindles in real time during sleep (Antony et al., 2018). Following a sleep spindle, there is a refractory period during which no new spindles occur. Reasoning that a lack of spindles would decrease the likelihood of reactivation and therefore consolidation, a real-time algorithm was used to present sound cues either immediately following spindles (in the refractory period) or after some time had passed (when spindles could occur again). Spatial memory was better for items cued outside the refractory period, further confirming the role of sleep spindles in effective memory consolidation (Antony et al., 2018).

4 Retrieval

Once new information has been encoded into memory and consolidated, the resultant memory trace can be brought back to mind, either spontaneously as a result of a relevant cue in the environment or intentionally as a result of effortful search. This process of remembering information that has been stored in long-term memory is called retrieval. Successful retrieval involves recapitulating the brain states that were present at

encoding. Not surprisingly, and much like encoding and consolidation, memory retrieval is driven by multiple brain circuits distributed across brain regions.

4.1 Neural Basis and Measures of Retrieval

When reviewing encoding, we highlighted brain regions that are more active for information that is subsequently remembered vs. forgotten in a later memory test. The neural basis of retrieval has also been studied by sorting brain activity in the memory test itself according to whether each old item is successfully remembered vs. forgotten (e.g., Buckner et al., 1998). The regions identified by this retrieval success contrast overlap with the regions involved in encoding (see Rugg and Vilberg, 2013; Spaniol et al., 2009). However, the relative activation in these regions differs: left ventrolateral PFC and MTL regions are more strongly involved in encoding, whereas left superior parietal and dorsolateral and anterior PFC regions are more strongly involved in retrieval. The proposed role of the PFC in retrieval is top-down selection of information, updating of episodic features, and guiding behavioral responses to the task at hand (Eichenbaum, 2000; Rugg et al., 2002). The proposed role of the parietal cortex (in particular, the angular gyrus) is guiding search or attention processes (Cabeza et al., 2012) and binding episodic details (Rugg et al., 2002). In addition to greater activity, the functional connectivity between these regions is higher during successful retrieval and the extent of increased connectivity in this network correlates with individual differences in recollection (King et al., 2015).

Successful retrieval is also associated with the reinstatement of brain states from encoding (see Frankland et al., 2019; Rissman and Wagner, 2012). MVPA studies have revealed that fine-grained spatial patterns of activity corresponding to the encoded content are recapitulated during retrieval (Johnson et al., 2009; Kuhl et al., 2011; Polyn et al., 2005). In one study (Polyn et al., 2005), participants encoded three lists consisting of either faces, places, or objects. During retrieval, participants were asked to freely recall items in any order and regardless of category. Nevertheless, the category of recalled items could be predicted in advance based on whole-brain patterns of activity. That is, returning the brain to a category-specific brain state from encoding facilitated retrieval of items from that category.

The degree to which an encoding brain state is reinstated correlates with memory performance. In one study (Kuhl et al., 2011), participants encoded two lists of word-image pairs. Some words were paired with just one image, whereas others were paired with a face in one list and a scene in the other list. During retrieval, participants were cued with a word and instructed to recall the image that was most recently paired with that word. A classifier was trained to distinguish faces and scenes based on activity patterns in the ventral occipitotemporal cortex. Performance was better when participants were able to recall the specific image compared to when they were only able to recall the category. This suggests a relationship between the fidelity of neural reinstatement and retrieval success and specificity. This relationship has since been reported in other modality-specific regions, as well as in frontal and parietal cortex, and is modulated by the hippocampus (Bosch et al., 2014; Danker et al., 2017; Ritchey et al., 2013).

In addition to retrieval success, an “all-or-none” memory measure, several studies have shown that neural reinstatement predicts rich aspects of memory performance during retrieval, such as the level of detail (Wing et al., 2015), vividness (Bone et al., 2020; Johnson et al., 2015; St-Laurent et al., 2015), and confidence (Thakral et al., 2015). These different aspects of memory retrieval seem to be supported by distinct neural mechanisms (Richter et al., 2016), revealing potential targets for memory enhancement through neurofeedback.

Complementing these findings from the fMRI literature, studies using iEEG have tracked the temporal dynamics associated with retrieval at millisecond resolution. Successful retrieval is associated with greater connectivity among various regions including the MTL, PFC, and parietal cortex (Watrous et al., 2013). In a more recent study (Vaz et al., 2020), single-unit spikes and local field potentials were recorded while participants performed a memory task. Ripple oscillations in the temporal cortex were found to reflect bursts of single-unit spiking activity organized into memory-specific sequences. These sequences occurred repeatedly during memory formation. Most intriguingly, these sequences were also replayed during successful retrieval and the extent of this replay in temporal cortex was associated with ripples in MTL. These results demonstrate that memory retrieval is supported by a reinstatement of a temporally stereotyped pattern of neural activity that was present during encoding.

Together, these findings suggest that successful memory retrieval is accompanied by coordinated activity between memory systems and frontoparietal cortex, often recapitulating spatiotemporal patterns of activity

that were present during initial encoding. However, due to the correlational nature of the methods used in these studies, they cannot pinpoint the causal roles of the different brain regions or rule out the possibility that some of the correlations observed between brain states and retrieval (e.g., reinstatement) may be epiphenomenal. Using rtfMRI, participants could be trained to target or elicit activity patterns in retrieval-related brain regions. By observing the effects on performance, the relative roles of these regions and the directionality of their interactions could be inferred. To a more practical end, neurofeedback could be used to impose beneficial retrieval states.

4.2 Methodological Approaches for Perturbing Retrieval

The idea that perturbation of the brain can prompt memory retrieval has a long history, going back to pioneering studies in which brain stimulation caused *deja vu*, a memory phenomenon akin to familiarity (Penfield & Perot, 1963). This direct link between induced brain activity and memory retrieval has been replicated more systematically (Jacobs et al., 2012). Stimulation at sites throughout the temporal lobe is associated with phenomena resembling memory retrieval (see Selimbeyoglu and Parvizi, 2010). In right dorsolateral PFC, excitatory TMS during retrieval leads to faster memory response times, suggestive of increased retrieval efficiency, whereas inhibitory TMS, reduces memory accuracy (Gagnon et al., 2011). In precuneus, TMS selectively improves source memory (Bonni et al., 2015). In addition to the memory benefit observed in healthy young adult population, repetitive TMS applied to the precuneus has been shown to improve retrieval-related symptoms of Alzheimer’s disease (Koch et al., 2018).

Rather than focusing on a single site, one recent study targeted multiple brain regions simultaneously (Kim et al., 2018). Patients navigated a virtual city to visit five stores. Later they had to identify where the stores were located (i.e., spatial memory) and the order in which they visited them (i.e., temporal memory). The functional connectivity between left inferior temporal sulcus and left middle frontal gyrus was selectively associated with spatial memory but not temporal memory. In a clear demonstration of functional specificity, stimulation of these network hubs just prior to retrieval impaired spatial but not temporal memory.

We have thus far discussed memory retrieval as it is studied in the laboratory. Participants are presented with a set of items and are later tested to see which ones are remembered. Another form of memory retrieval involves recapitulating one’s own autobiographical experiences when prompted (e.g., Benoit and Schacter, 2015). This process too can be influenced by direct manipulation of brain states. In a recent study, researchers applied TMS to angular gyrus during an autobiographical recall task. When participants were asked to recall an event from the last five years of their life, memories recounted during TMS had fewer episodic details (Thakral et al., 2017). This provides evidence for the necessity of the angular gyrus in retrieving specific contextual details from memory.

Leveraging the idea that reinstatement of relevant brain states drives successful retrieval, a recent study used closed-loop rtfMRI to directly test whether context reinstatement leads to better memory retrieval (deBettencourt et al., 2019). In this study (Figure 2A), two contexts were created by interspersing images of either scenes or faces between words from two different lists during encoding. Prior to recall, the participants were instructed to reinstate the context from one of the lists. They were provided with a visual neurofeedback display, which was a weighted blend of a face and a scene. Using a classifier trained to distinguish brain activity patterns for faces vs. scenes, the visibility of the cued context in the display tracked the degree of classifier evidence for the cued context in the brain. For example, if cued to reinstate the context of the list interspersed with scenes, the more the participant’s brain contained information about scenes, the more strongly the scene appeared in the face/scene blend. This created a positive feedback loop that amplified subtle inflections in brain state. Importantly, the cued context was not always a valid indicator of which list would ultimately be tested. For example, sometimes participants were cued to reinstate a scene context, but then told to recall the words from the list with the face context. Higher levels of neural context reinstatement (operationalized in terms of classifier evidence for the cued context) were associated with better recall if the test list matched the reinstated context and worse recall if the participants were asked to reinstate the wrong context (Figure 2B). Critically, this relationship between neural context reinstatement and free recall was only detectable with neurofeedback; it was eliminated in a control group who were presented with fully coherent images of a category as contextual retrieval cues irrespective of brain state (Figure 2C).

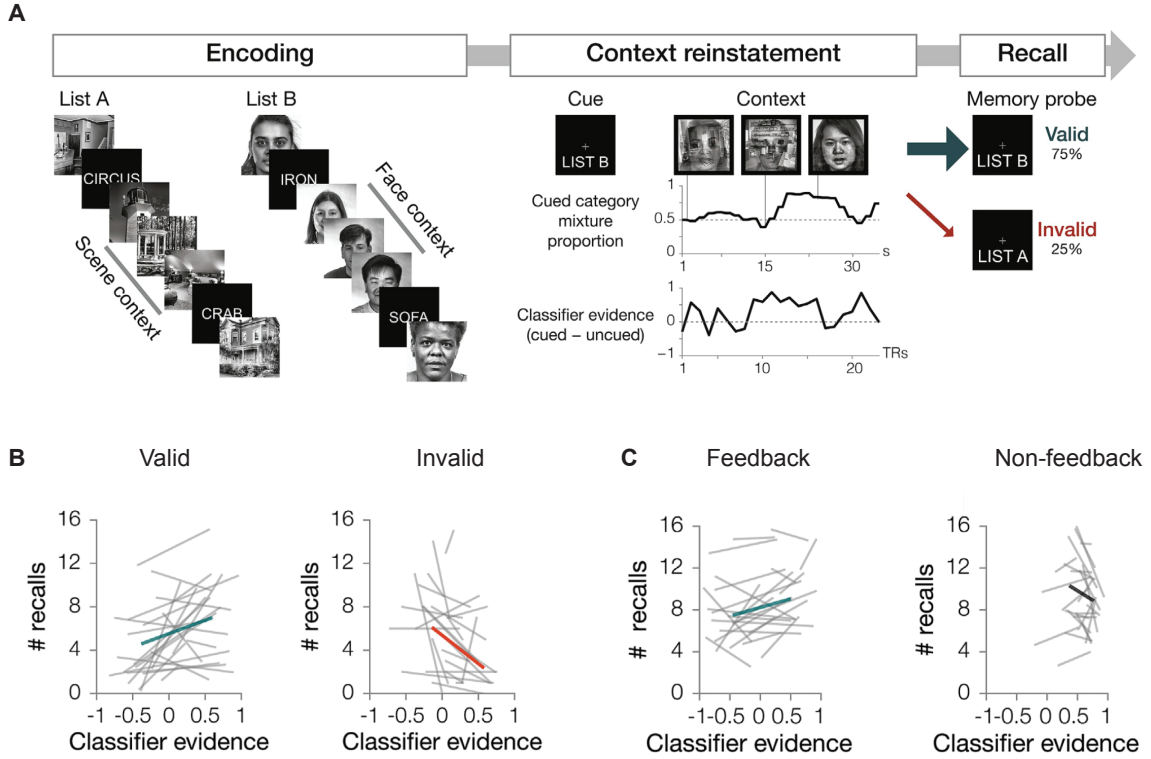


Figure 2: Neurofeedback helps to reveal a relationship between context reinstatement and memory retrieval (deBettencourt et al., 2019). (A) Participants first encoded two lists of words paired with a visual context of scenes or faces. They were then cued to reinstate one of the two contexts during neurofeedback with the aid of composite face/scene images. The mixture proportion of the two categories in the composite image was determined by the classifier evidence for the cued relative to uncued category context, based on real-time decoding of incoming brain images. Finally, participants were cued to recall words from one of the two lists, which could either be valid (matching the context cued for reinstatement during neurofeedback) or invalid (mismatching the reinstated context). (B) Context reinstatement as measured by classifier evidence benefited the recall of words when valid (left) but hurt recall when invalid (right). (C) The relationship between neural context reinstatement and memory recall performance was only detectable with neurofeedback (left). When the image was set to always be fully coherent (i.e., containing only the cued category), and thus decoupled from neural reinstatement, there was no relationship between neural reinstatement and memory performance (right). Figure adapted with permission.

5 Studying Memory using Real-Time fMRI

fMRI has proven useful for understanding memory because it provides whole-brain coverage, can be used noninvasively in the healthy brain, and allows neural data to be collected concurrently with all three stages of memory reviewed above. rtfMRI confers additional advantages for memory research, complementing conventional fMRI by establishing and manipulating brain-behavior relationships and testing causal hypotheses about memory processes. Given that memory is supported by interactions across widely distributed brain regions, especially deep-brain structures such as the hippocampus and MTL cortex, rtfMRI provides advantages relative to other tools (e.g., TMS or DES) that do not provide access to the whole brain. With recent developments in MRI technology and computational tools for fMRI data analysis (Cohen et al., 2017), rtfMRI

provides alluring new opportunities to study the spatial and temporal dynamics of the neural mechanisms underlying memory.

The most defining characteristic of rtfMRI is that the incoming volumes of whole-brain data are analyzed on the fly (often in under a few seconds) and the results are used to modify the ongoing experiment. Paradigms using these methods typically fall into two broad categories: triggering and neurofeedback. In a triggering design (Figure 1B), the goal is to determine the timing and content of stimulus presentations based on the ongoing brain states of participants. Volumes are analyzed as they are collected and compared to some researcher-defined benchmark. Then researchers simply wait for the desired state, and when it is detected, trigger a particular type of trial, such as a new item to encode (Yoo et al., 2012).

In a neurofeedback design (Figure 1C), the primary goal is to induce a target brain state (e.g., activation in a certain brain region or a particular spatial pattern) in the participants. As each incoming volume is collected and analyzed, feedback is given based on some measure of the similarity of the current brain state to the desired brain state. The most common way of providing feedback is via a visual signal that scales with the similarity to the target brain state (e.g., a looming circle; Shibata et al., 2011). In this case, the underlying mechanism can be seen as a version of operant conditioning where participants learn to self-regulate brain activity instead of behavior, via repeated trial-and-error, to maximize the feedback signal. More recently, some neurofeedback studies have provided feedback signals that are more integrated into the experiment, often referred to as closed-loop (deBettencourt et al., 2015; deBettencourt et al., 2019; Mennen et al., 2020). For example, similarity to the target brain state may be used to control task difficulty (deBettencourt et al., 2019), such that successful regulation of brain activity is rewarded with an easier task. Both triggering and neurofeedback paradigms can be based on univariate activation in one region, the functional connectivity between regions, a multivariate pattern of activity across voxels, or even a combination of these measures. As a result of the flexibility of this method, there are many opportunities for using rtfMRI to study the mechanics underlying memory. Below, we highlight just a few.

5.1 Enhancing Encoding or Retrieval with Real-time fMRI

Successful memory depends not just on the activation of one single region but rather, on the interaction and coordination of distributed brain networks. Recently, neurofeedback has been applied to enhance functional networks relevant for behavior and yielded some promising results (e.g., Megumi et al., 2015; Ramot et al., 2017; Scheinost et al., 2020; Yamashita et al., 2017). A potential application of rtfMRI is to use neurofeedback to train participants to induce optimal encoding or retrieval states by targeting the functional networks identified in previous studies as being critical to these processes. In this way, participants can be taught how to attain the optimal encoding or retrieval brain states, so that they can learn to regulate their own functional networks to improve performance going forward.

In addition to being a potential novel tool for memory enhancement, using neurofeedback to target encoding vs. retrieval networks could also answer important theoretical questions that are not easily addressed with conventional neuroimaging methods. Encoding and retrieval involve overlapping brain regions in the MTL cortex, hippocampus, PFC, and parietal cortex (see Rugg and Vilberg, 2013; Spaniol et al., 2009). Using neurofeedback to target these regions during encoding and retrieval may shed light on the different roles they play in these processes. For instance, we know which regions are implicated in encoding, and have identified sets of regions whose connectivity is associated with better subsequent memory. However, using traditional methods it is difficult to pinpoint which regions are the most pivotal in driving encoding processes, or the directionality of these regional interactions. Recent advances in computational modeling of network dynamics suggest that some brain regions will have a much broader influence over global brain states than others (Bassett & Khambhati, 2017; Muldoon et al., 2016). Assuming that this principle extends to the networks that drive encoding as well, rtfMRI could be used to identify and subsequently target the regions most likely to propagate changes in connectivity across the entire network. Specifically, individual nodes in the encoding network could be systematically upregulated using neurofeedback, to determine which regions most effectively propagate changes in the activity or connectivity of the entire network. The insights gained from this approach could lead to the development of more targeted intervention procedures for memory enhancement.

5.2 Inducing or Disrupting Item or Category-Specific Representations

Hippocampal replay during consolidation of learned information improves later memory (see Foster, 2017; Tambini and Davachi, 2019). As discussed above, EEG can be used to detect sleep stages and spindles in real-time, and both types of information can be used to optimize the timing of targeted memory reactivation cues. rtfMRI has complementary strengths and weaknesses from EEG with regard to studying and promoting consolidation. Due to its worse temporal resolution, fMRI is less useful than EEG for sleep staging and spindle tracking. On the positive side, the superior spatial resolution of fMRI can potentially provide a more specific way of monitoring or instantiating replay in one or several regions. The spatial resolution of fMRI is also advantageous when it comes to tracking the specific content of the replayed information. Although there have been some basic successes in decoding the content of replayed memories during sleep with EEG (faces vs. scenes: Cairney et al., 2018; Schönauer et al., 2017; left- vs. right-side movements: Wang et al., 2019), the spatial blurring of EEG places an upper bound on the resolution of decoding. By contrast, fMRI has been used to decode replay of more fine-grained content (e.g., specific images of satellites within categories: Schapiro et al., 2018).

Using rtfMRI, researchers could apply both triggering and neurofeedback approaches to study memory consolidation. Scans during an initial encoding phase could be used to quantify the spatial patterns of brain activity associated with exposure to a particular item, category of items, or context. Next, while participants are awake but at rest in a subsequent phase, researchers could extract information about the participant’s internal representations by analyzing multi-voxel patterns from each incoming volume. By monitoring the moment-to-moment fluctuations in internal states, activity consistent with the prescribed stimulus or category could be detected and used to precisely time an experimental intervention (Figure 1B). For example, if trying to prevent replay of a particular face, activation of that face representation, detected in real time, could be met with distracting sounds, which have been shown to disrupt replay (Schreiner et al., 2015). Further, the target brain region and timing of interference could be manipulated to establish the causal roles of brain regions in replay during consolidation.

Although triggering approaches could help reveal mechanisms of consolidation by disrupting replay, neurofeedback approaches might have more potential for real-world application. Namely, participants could be trained to become better at replaying or to learn mental strategies that promote replay. That is, during a post-encoding period, researchers could display a continuously changing feedback signal to a participant, reflecting classifier evidence in favor of a target item or category whose memory they are trying to enhance. By selectively targeting the hippocampus or different neocortical regions, this approach could help disentangle the relative contributions of hippocampal vs. cortical replay to later memory, and help establish the direction of information flow between the hippocampus and cortex (see Rothschild et al., 2017).

5.3 Testing Theories of Learning-Related Change

As we learn, representations in memory are constantly reorganized. Memories are weakened and strengthened, but can also become more or less similar to one another. One theory of neuroplasticity, the *nonmonotonic plasticity hypothesis* (NMPH; Ritvo et al., 2019) proposes that these changes, both in individual memories and in the relationships between them, are driven by varying degrees of concurrent activation across memories (Detre et al., 2013; Hulbert & Norman, 2015; Newman & Norman, 2010). According to the NMPH, the degree of shared activation among two memory representations will determine whether they integrate or differentiate: While strong co-activation will lead to integration of two memories, moderate co-activation will lead to their differentiation, and lower levels of co-activation will lead to no change in the similarity structure of the memories.

It is challenging to test this hypothesis using traditional fMRI methods, because there is no way of knowing *a priori* where a particular experimental condition will fall on the continuum from low co-activation to moderate co-activation (leading to differentiation) to strong co-activation (leading to integration). This is further complicated by the fact that co-activation levels can vary across individuals and across trials within a particular experimental condition. One approach to this problem is to include a variety of conditions meant to evoke different levels of activity, measure those levels of activity *post-hoc* using MVPA, and then relate measured activity to representational change, all the while hoping that enough observations are collected at each level of co-activation to test the predictions of the NMPH. rtfMRI could provide a more efficient way of testing these predictions: Specifically, one could estimate the neural representations of a pair of items (A and

B) within each individual. Then, during presentation of the B item, neurofeedback could be used to reward participants for activating the A item to a particular "target" level, without actually displaying A to the participant. For example, if participants were rewarded for strong A activation during presentation of B, this should induce strong co-activation of A and B, leading to integration of the two memories (Figure 1C). More generally, this approach provides an exciting opportunity to establish a causal relationship between particular levels of co-activation of memories, representational changes in the brain, and behavioral effects of these changes.

5.4 Leveraging Expert Brain States to Enhance Learning

Memory in the real world often deals with complex, abstract concepts that are learned gradually and intermixed with other information. Moreover, there are dramatic differences in how well people learn or remember when presented with the same content (Jonassen & Grabowski, 2012). What underlying neural mechanisms could explain such individual differences? Furthermore, could we train individuals to become better at learning a certain topic by teaching them how to modulate their own brain activity?

Previous studies have found that comprehension and memory of narratives can be predicted from the similarity between the neural patterns in listeners and that of the speaker (Hasson et al., 2012; Nguyen et al., 2019; Stephens et al., 2010). The efficacy of learning could thus be predicted by the neural similarity between a teacher/expert and students (Meshulam et al., 2020). This could be leveraged by rtfMRI studies that directly reward similarity to a teacher's neural patterns and measure the effect on later comprehension and learning of the presented content. Specifically, one could measure the dynamic brain states present in a set of teachers or experts while they are outlining a particularly difficult concept. As a student is learning, their dynamic brain states could be analyzed in real time, and feedback given to nudge their brains toward (or away from) states that more closely resemble those observed in the experts. Researchers could then probe for differences in performance as a result of this neurofeedback manipulation. In addition to attempting to induce appropriate brain states for learning, the similarity of brain states could be used to provide a sensitive measure of learning in real time, especially in cases where the strategy for improvement might not be easily verbalized. Techniques like these have value for basic science by establishing causal links between changes in brain states and the efficacy of learning, but could also eventually find application in pedagogical settings (perhaps with different neural technologies) for brain-based education and evaluation.

6 Conclusion

Memory is fundamental to every aspect of the mind and brain, and ultimately may underlie all rtfMRI training effects described in this book. Our focus in this chapter, however, was on what could be learned in the other direction, using rtfMRI to understand memory itself. Memory processes are underpinned by dynamic processes in the hippocampus and MTL cortex, and more broadly throughout cortex. The enterprise of using rtfMRI to study these processes is still in its infancy, relative to other clinical and basic cognitive approaches. Several other methods (e.g., DES, TMS, EEG) have been used to monitor or manipulate these neural circuits and shed light on how they support memory, albeit with mixed results. We propose that rtfMRI, through broad sampling of the whole brain with relatively high spatial sensitivity (especially to deep-brain memory structures), will allow for more flexible control over the neural processes driving encoding, consolidation, and retrieval. Complementing other methods, rtfMRI provides unprecedented opportunities to monitor, perturb, and instantiate brain states, in order to draw causal inferences about how brain dynamics support optimal memory performance and guide the development of memory enhancement strategies.

References

- Antony, J. W., Piloto, L., Wang, M., Pacheco, P., Norman, K. A., & Paller, K. A. (2018). Sleep spindle refractoriness segregates periods of memory reactivation. *Current Biology*, 28(11), 1736–1743.
- Bassett, D. S., & Khambhati, A. N. (2017). A network engineering perspective on probing and perturbing cognition with neurofeedback. *Annals of the New York Academy of Sciences*, 1396(1), 126–143.
- Benoit, R. G., & Schacter, D. L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia*, 75(1), 450–457.
- Bone, M. B., Ahmad, F., & Buchsbaum, B. R. (2020). Feature-specific neural reactivation during episodic memory. *Nature Communications*, 11(1), 1–13.
- Bonni, S., Veniero, D., Mastropasqua, C., Ponzo, V., Caltagirone, C., Bozzali, M., & Koch, G. (2015). TMS evidence for a selective role of the precuneus in source memory retrieval. *Behavioural Brain Research*, 282, 70–75.
- Bosch, S. E., Jehee, J. F., Fernández, G., & Doeller, C. F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *Journal of Neuroscience*, 34(22), 7493–7500.
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered. *Science*, 281(5380), 1185–1187.
- Buckner, R. L., Koutstaal, W., Schacter, D. L., Dale, A. M., Rotte, M., & Rosen, B. R. (1998). Functional-anatomic study of episodic retrieval: Ii. selective averaging of event-related fMRI trials to test the retrieval success hypothesis. *NeuroImage*, 7(3), 163–175.
- Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: An integrative theoretical account. *Trends in Cognitive Sciences*, 16(6), 338–352.
- Cairney, S. A., El Marj, N., Staresina, B. P., Et al. (2018). Memory consolidation is linked to spindle-mediated information processing during sleep. *Current Biology*, 28(6), 948–954.
- Cohen, J. D., Daw, N., Engelhardt, B., Hasson, U., Li, K., Niv, Y., Norman, K. A., Pillow, J., Ramadge, P. J., Turk-Browne, N. B., Et al. (2017). Computational approaches to fMRI analysis. *Nature Neuroscience*, 20(3), 304–313.
- Corkin, S. (2013). *Permanent present tense: The unforgettable life of the amnesic patient, HM*. Basic Books (AZ).
- Danker, J. F., Tompary, A., & Davachi, L. (2017). Trial-by-trial hippocampal encoding activation predicts the fidelity of cortical reinstatement during subsequent retrieval. *Cerebral Cortex*, 27(7), 3515–3524.
- Daselaar, S. M., Prince, S. E., Dennis, N. A., Hayes, S. M., Kim, H., & Cabeza, R. (2009). Posterior midline and ventral parietal activity is associated with retrieval success and encoding failure. *Frontiers in Human Neuroscience*, 3, 13.
- deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-browne, N. B. (2015). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience*, 18(3), 470–475.
- deBettencourt, M. T., Turk-Browne, N. B., & Norman, K. A. (2019). Neurofeedback helps to reveal a relationship between context reinstatement and memory retrieval. *NeuroImage*, 200, 292–301.
- Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, 51(12), 2371–2388.
- Dudai, Y. (2012). The restless engram: Consolidations never end. *Annual Review of Neuroscience*, 35, 227–247.
- Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience*, 1(1), 41–50.
- Ezzyat, Y., Kragel, J. E., Burke, J. F., Levy, D. F., Lyalenko, A., Wanda, P., O’Sullivan, L., Hurley, K. B., Busygin, S., Pedisich, I., Et al. (2017). Direct brain stimulation modulates encoding states and memory performance in humans. *Current Biology*, 27(9), 1251–1258.
- Ezzyat, Y., Wanda, P. A., Levy, D. F., Kadel, A., Aka, A., Pedisich, I., Sperling, M. R., Sharan, A. D., Lega, B. C., Burks, A., Et al. (2018). Closed-loop stimulation of temporal cortex rescues functional networks and improves memory. *Nature Communications*, 9(1), 1–8.
- Foster, D. J. (2017). Replay comes of age. *Annual Review of Neuroscience*, 40, 581–602.
- Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, 6(2), 119–130.
- Frankland, P. W., Josselyn, S. A., & Köhler, S. (2019). The neurobiological foundation of memory retrieval. *Nature Neuroscience*, 22(10), 1576–1585.

- Gagnon, G., Blanchet, S., Grondin, S., & Schneider, C. (2010). Paired-pulse transcranial magnetic stimulation over the dorsolateral prefrontal cortex interferes with episodic encoding and retrieval for both verbal and non-verbal materials. *Brain Research*, 1344, 148–158.
- Gagnon, G., Schneider, C., Grondin, S., & Blanchet, S. (2011). Enhancement of episodic memory in young and healthy adults: A paired-pulse TMS study on encoding and retrieval performance. *Neuroscience Letters*, 488(2), 138–142.
- Goyal, A., Miller, J., Watrous, A. J., Lee, S. A., Coffey, T., Sperling, M. R., Sharan, A., Worrell, G., Berry, B., Lega, B., Et al. (2018). Electrical stimulation in hippocampus and entorhinal cortex impairs spatial and temporal memory. *Journal of Neuroscience*, 38(19), 4471–4481.
- Guderian, S., Schott, B. H., Richardson-Klavehn, A., & Düzel, E. (2009). Medial temporal theta state before an event predicts episodic encoding success in humans. *Proceedings of the National Academy of Sciences*, 106(13), 5365–5370.
- Hanlon, C. (2017). Blunt or precise? a note about the relative precision of figure-of-eight rTMS coils. *Brain Stimulation*, 10(2), 338–339.
- Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, 16(2), 114–121.
- Hermiller, M. S., Chen, Y. F., Parrish, T. B., & Voss, J. L. (2020). Evidence for immediate enhancement of hippocampal memory encoding by network-targeted theta-burst stimulation during concurrent fmri. *Journal of Neuroscience*.
- Hulbert, J., & Norman, K. A. (2015). Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cerebral Cortex*, 25(10), 3994–4008.
- Inman, C. S., Manns, J. R., Bijanki, K. R., Bass, D. I., Hamann, S., Drane, D. L., Fasano, R. E., Kovach, C. K., Gross, R. E., & Willie, J. T. (2018). Direct electrical stimulation of the amygdala enhances declarative memory in humans. *Proceedings of the National Academy of Sciences*, 115(1), 98–103.
- Jacobs, J., Lega, B., & Anderson, C. (2012). Explaining how brain stimulation can evoke memories. *Journal of Cognitive Neuroscience*, 24(3), 553–563.
- Jacobs, J., Miller, J., Lee, S. A., Coffey, T., Watrous, A. J., Sperling, M. R., Sharan, A., Worrell, G., Berry, B., Lega, B., Et al. (2016). Direct electrical stimulation of the human entorhinal region and hippocampus impairs memory. *Neuron*, 92(5), 983–990.
- Johnson, J. D., McDuff, S. G., Rugg, M. D., & Norman, K. A. (2009). Recollection, familiarity, and cortical reinstatement: A multivoxel pattern analysis. *Neuron*, 63(5), 697–708.
- Johnson, M. K., Kuhl, B. A., Mitchell, K. J., Ankudowich, E., & Durbin, K. A. (2015). Age-related differences in the neural basis of the subjective vividness of memories: Evidence from multivoxel pattern classification. *Cognitive, Affective, & Behavioral Neuroscience*, 15(3), 644–661.
- Jonassen, D. H., & Grabowski, B. L. (2012). *Handbook of individual differences, learning, and instruction*. Routledge.
- Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., & Buckner, R. L. (2008). Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, 100(1), 129–139.
- Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage*, 54(3), 2446–2461.
- Kim, K., Schedlbauer, A., Rollo, M., Karunakaran, S., Ekstrom, A. D., & Tandon, N. (2018). Network-based brain stimulation selectively impairs spatial retrieval. *Brain Stimulation*, 11(1), 213–221.
- King, D. R., de Chastelaine, M., Elward, R. L., Wang, T. H., & Rugg, M. D. (2015). Recollection-related increases in functional connectivity predict individual differences in memory accuracy. *Journal of Neuroscience*, 35(4), 1763–1772.
- Koch, G., Bonni, S., Pellicciari, M. C., Casula, E. P., Mancini, M., Esposito, R., Ponzo, V., Picazio, S., Di Lorenzo, F., Serra, L., Et al. (2018). Transcranial magnetic stimulation of the precuneus enhances memory and neural activity in prodromal alzheimer’s disease. *NeuroImage*, 169, 302–311.
- Kucewicz, M. T., Berry, B. M., Miller, L. R., Khadjevand, F., Ezzyat, Y., Stein, J. M., Kremen, V., Brinkmann, B. H., Wanda, P., Sperling, M. R., Et al. (2018). Evidence for verbal memory enhancement with electrical brain stimulation in the lateral temporal cortex. *Brain*, 141(4), 971–978.
- Kuhl, B. A., Rissman, J., Chun, M. M., & Wagner, A. D. (2011). Fidelity of neural reactivation reveals competition between memories. *Proceedings of the National Academy of Sciences*, 108(14), 5903–5908.

- Kuhl, B. A., Rissman, J., & Wagner, A. D. (2012). Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia*, 50(4), 458–469.
- Lewis-Peacock, J. A., & Norman, K. A. (2014). Multi-voxel pattern analysis of fMRI data. *The Cognitive Neurosciences*, 512, 911–920.
- Libby, L. A., Ekstrom, A. D., Ragland, J. D., & Ranganath, C. (2012). Differential connectivity of perirhinal and parahippocampal cortices within human hippocampal subregions revealed by high-resolution functional imaging. *Journal of Neuroscience*, 32(19), 6550–6560.
- Marr, D., Willshaw, D., & McNaughton, B. (1991). Simple memory: A theory for archicortex, In *From the retina to the neocortex*. Springer.
- 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.
- McGaugh, J. L. (1966). Time-dependent processes in memory storage. *Science*, 153(3742), 1351–1358.
- Megumi, F., Yamashita, A., Kawato, M., & Imamizu, H. (2015). Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. *Frontiers in Human Neuroscience*, 9, 160.
- Mennen, A. C., Turk-Browne, N. B., Wallace, G., Seok, D., Jaganjac, A., Stock, J., deBettencourt, M. T., Cohen, J. D., Norman, K. A., & Sheline, Y. I. (2020). Closed-loop fMRI neurofeedback to reduce negative attentional bias in depression. *bioRxiv*.
- Merkow, M. B., Burke, J. F., Ramayya, A. G., Sharan, A. D., Sperling, M. R., & Kahana, M. J. (2017). Stimulation of the human medial temporal lobe between learning and recall selectively enhances forgetting. *Brain Stimulation*, 10(3), 645–650.
- Meshulam, M., Hasenfratz, L., Hillman, H., Liu, Y.-F., Nguyen, M., Norman, K. A., & Hasson, U. (2020). Think like an expert: Neural alignment predicts understanding in students taking an introduction to computer science course. *bioRxiv*.
- Misanin, J. R., Miller, R. R., & Lewis, D. J. (1968). Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. *Science*, 160(3827), 554–555.
- Muldoon, S. F., Pasqualetti, F., Gu, S., Cieslak, M., Grafton, S. T., Vettel, J. M., & Bassett, D. S. (2016). Stimulation-based control of dynamic brain networks. *PLoS Computational Biology*, 12(9), e1005076.
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, 7(2), 217–227.
- Nader, K., & Hardt, O. (2009). A single standard for memory: The case for reconsolidation. *Nature Reviews Neuroscience*, 10(3), 224–234.
- Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, 406(6797), 722–726.
- Newman, E. L., & Norman, K. A. (2010). Moderate excitation leads to weakening of perceptual representations. *Cerebral Cortex*, 20(11), 2760–2770.
- Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage*, 184, 161–170.
- Paller, K. A., & Norman, K. A. (2020). Replay-based consolidation governs enduring memory storage. *The Cognitive Neurosciences*, 265–276.
- Park, H., & Rugg, M. D. (2010). Prestimulus hippocampal activity predicts later recollection. *Hippocampus*, 20(1), 24–28.
- Penfield, W., & Perot, P. (1963). The brain's record of auditory and visual experience: A final summary and discussion. *Brain*, 86(4), 595–696.
- Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science*, 310(5756), 1963–1966.
- Ramot, M., Kimmich, S., Gonzalez-Castillo, J., Roopchansingh, V., Popal, H., White, E., Gotts, S. J., & Martin, A. (2017). Direct modulation of aberrant brain network connectivity through real-time neurofeedback. *Elife*, 6, e28974.
- Ranganath, C., Heller, A., Cohen, M. X., Brozinsky, C. J., & Rissman, J. (2005). Functional connectivity with the hippocampus during successful memory formation. *Hippocampus*, 15(8), 997–1005.
- Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, 13(10), 713–726.

- Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science*, 315(5817), 1426–1429.
- Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *Elife*, 5, e18260.
- Rissman, J., & Wagner, A. D. (2012). Distributed representations in memory: Insights from functional brain imaging. *Annual Review of Psychology*, 63, 101–128.
- Ritchey, M., Wing, E. A., LaBar, K. S., & Cabeza, R. (2013). Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cerebral Cortex*, 23(12), 2818–2828.
- Ritvo, V. J., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic plasticity: How memory retrieval drives learning. *Trends in Cognitive Sciences*, 23(9), 726–742.
- Rothschild, G., Eban, E., & Frank, L. M. (2017). A cortical–hippocampal–cortical loop of information processing during memory consolidation. *Nature Neuroscience*, 20(2), 251–259.
- Rugg, M. D., Otten, L. J., & Henson, R. N. (2002). The neural basis of episodic memory: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1424), 1097–1110.
- Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current Opinion in Neurobiology*, 23(2), 255–260.
- Schapiro, A. C., McDevitt, E. A., Rogers, T. T., Mednick, S. C., & Norman, K. A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nature Communications*, 9(1), 1–11.
- Scheinost, D., Hsu, T. W., Avery, E. W., Hampson, M., Constable, R. T., Chun, M. M., & Rosenberg, M. D. (2020). Connectome-based neurofeedback: A pilot study to improve sustained attention. *NeuroImage*, 116684.
- Schiller, D., Monfils, M.-H., Raio, C. M., Johnson, D. C., LeDoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, 463(7277), 49–53.
- Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, 1, 1–8.
- Schönauer, M., Alizadeh, S., Jamalabadi, H., Abraham, A., Pawlizki, A., & Gais, S. (2017). Decoding material-specific memory reprocessing during sleep in humans. *Nature Communications*, 8(1), 1–9.
- Schreiner, T., Lehmann, M., & Rasch, B. (2015). Auditory feedback blocks memory benefits of cueing during sleep. *Nature Communications*, 6(1), 1–11.
- Selimbeyoglu, A., & Parvizi, J. (2010). Electrical stimulation of the human brain: Perceptual and behavioral phenomena reported in the old and new literature. *Frontiers in Human Neuroscience*, 4, 46.
- Shibata, K., Watanabe, T., Sasaki, Y., & Kawato, M. (2011). Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science*, 334(6061), 1413–1415.
- Spaniol, J., Davidson, P. S., Kim, A. S., Han, H., Moscovitch, M., & Grady, C. L. (2009). Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. *Neuropsychologia*, 47(8-9), 1765–1779.
- Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, 107(32), 14425–14430.
- St-Laurent, M., Abdi, H., & Buchsbaum, B. R. (2015). Distributed patterns of reactivation predict vividness of recollection. *Journal of Cognitive Neuroscience*, 27(10), 2000–2018.
- Summerfield, C., Greene, M., Wager, T., Egner, T., Hirsch, J., & Mangels, J. (2006). Neocortical connectivity during episodic memory formation. *PLoS Biology*, 4(5).
- Suthana, N., & Fried, I. (2014). Deep brain stimulation for enhancement of learning and memory. *NeuroImage*, 85, 996–1002.
- Suthana, N., Haneef, Z., Stern, J., Mukamel, R., Behnke, E., Knowlton, B., & Fried, I. (2012). Memory enhancement and deep-brain stimulation of the entorhinal area. *New England Journal of Medicine*, 366(6), 502–510.
- Tambini, A., & Davachi, L. (2019). Awake reactivation of prior experiences consolidates memories and biases cognition. *Trends in Cognitive Sciences*, 23(10), 876–890.
- Tambini, A., Nee, D. E., & D’Esposito, M. (2018). Hippocampal-targeted theta-burst stimulation enhances associative memory formation. *Journal of Cognitive Neuroscience*, 30(10), 1452–1472.

- Thakral, P. P., Madore, K. P., & Schacter, D. L. (2017). A role for the left angular gyrus in episodic simulation and memory. *Journal of Neuroscience*, 37(34), 8142–8149.
- Thakral, P. P., Wang, T. H., & Rugg, M. D. (2015). Cortical reinstatement and the confidence and accuracy of source memory. *NeuroImage*, 109, 118–129.
- Tulving, E. (1985). *Elements of episodic memory*. Oxford: Clarendon Press.
- Turk-Browne, N. B., Yi, D.-J., & Chun, M. M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron*, 49(6), 917–927.
- Uncapher, M. R., & Wagner, A. D. (2009). Posterior parietal cortex and episodic encoding: Insights from fMRI subsequent memory effects and dual-attention theory. *Neurobiology of Learning and Memory*, 91(2), 139–154.
- Vaz, A. P., Wittig, J. H., Inati, S. K., & Zaghoul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. *Science*, 367(6482), 1131–1134.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281(5380), 1188–1191.
- Wang, B., Antony, J. W., Lurie, S., Brooks, P. P., Paller, K. A., & Norman, K. A. (2019). Targeted memory reactivation during sleep elicits neural signals related to learning content. *Journal of Neuroscience*, 39(34), 6728–6736.
- Wang, J. X., Rogers, L. M., Gross, E. Z., Ryals, A. J., Dokucu, M. E., Brandstatt, K. L., Hermiller, M. S., & Voss, J. L. (2014). Targeted enhancement of cortical-hippocampal brain networks and associative memory. *Science*, 345(6200), 1054–1057.
- Ward, E. J., Chun, M. M., & Kuhl, B. A. (2013). Repetition suppression and multi-voxel pattern similarity differentially track implicit and explicit visual memory. *Journal of Neuroscience*, 33(37), 14749–14757.
- Watrous, A. J., Tandon, N., Conner, C. R., Pieters, T., & Ekstrom, A. D. (2013). Frequency-specific network connectivity increases underlie accurate spatiotemporal memory retrieval. *Nature Neuroscience*, 16(3), 349–356.
- Wing, E. A., Ritchey, M., & Cabeza, R. (2015). Reinstatement of individual past events revealed by the similarity of distributed activation patterns during encoding and retrieval. *Journal of Cognitive Neuroscience*, 27(4), 679–691.
- Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010). Greater neural pattern similarity across repetitions is associated with better memory. *Science*, 330(6000), 97–101.
- Yamashita, A., Hayasaka, S., Kawato, M., & Imamizu, H. (2017). Connectivity neurofeedback training can differentially change functional connectivity and cognitive performance. *Cerebral Cortex*, 27(10), 4960–4970.
- Yoo, J. J., Hinds, O., Ofen, N., Thompson, T. W., Whitfield-Gabrieli, S., Triantafyllou, C., & Gabrieli, J. D. (2012). When the brain is prepared to learn: Enhancing human learning using real-time fMRI. *NeuroImage*, 59(1), 846–852.