Trends in Neurosciences

CellPress REVIEWS

Opinion

Transforming the Concept of Memory Reactivation

Serra E. Favila,[1] Hongmi Lee,[2] and Brice A. Kuhl[3,*]

Reactivation refers to the phenomenon wherein patterns of neural activity expressed during perceptual experience are re-expressed at a later time, a putative neural marker of memory. Reactivation of perceptual content has been observed across many cortical areas and correlates with objective and subjective expressions of memory in humans. However, because reactivation emphasizes similarities between perceptual and memory-based representations, it obscures differences in how perceptual events and memories are represented. Here, we highlight recent evidence of systematic differences in how (and where) perceptual events and memories are represented in the brain. We argue that neural representations of memories are best thought of as spatially transformed versions of perceptual representations. We consider why spatial transformations occur and identify critical questions for future research.

Moving beyond Memory Reactivation

When remembering an event from the past, it often feels as though we are re-experiencing the content of that event, bringing to mind the sensations, emotions, or thoughts that characterized the initial experience. This subjective quality of memory is paralleled by an important neural phenomenon: that **content-sensitive patterns of neural activity** (see Glossary) evoked during the initial experience of an event are re-expressed when that event is retrieved from memory. This phenomenon is referred to as **reactivation** and has been an influential tool for studying **episodic memory** in humans. Neural measures of reactivation have been shown to be predictive of the accuracy [1–3], subjective vividness [4,5], and consequences of **memory retrieval** [6–8]. Another notable aspect of reactivation is that it can be measured in contexts such as rest [9–11] or sleep [12,13] when behavioral recordings are not feasible. Over the past two decades, there have been tremendous advances in both the computational sophistication and the sensitivity of the methods used to measure reactivation (Box 1), but these methods exploit the same fundamental phenomenon of a ‘match’ between content-sensitive neural activity patterns evoked during perception and memory retrieval. While this emphasis on the similarity between perceptual and memory-based representations is appealing and is unquestionably important, here we argue that it also provides an incomplete view of how the contents of memories are expressed in the brain during memory retrieval. In particular, because reactivation, by definition, tests for similarities between perception and retrieval, it necessarily fails to capture ways in which content representations differ between perception and memory.

One obvious way in which memory-based and perceptual representations might differ is in terms of the amount of information they contain, with memory-based representations presumably being a weaker or degraded version of perceptual representations. Empirically, this might manifest, for example, as a reduced ability to decode the content of an event from neural activity patterns evoked during retrieval compared with patterns evoked during perception. However, a number of recent studies have demonstrated clear violations of this idea: specifically, some brain regions

Highlights

A foundational finding in the field of memory is that content-sensitive patterns of neural activity expressed during perceptual experiences are re-expressed when experiences are remembered, a phenomenon termed reactivation. However, reactivation obscures key differences in how perceptual events and memories are represented in the brain.

Recent findings suggest systematic, spatial transformations of content-sensitive neural activity patterns from perception to memory retrieval. These transformations occur within sensory cortex and from sensory cortex to frontoparietal cortex.

We consider why spatial transformations occur and identify critical questions to be addressed in future research. Understanding the ways in which memory representations differ from perceptual representations will critically inform theoretical accounts of memory and will help clarify how the brain recreates the past.

[1]Department of Psychology, Columbia University, New York, NY 10027, USA
[2]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218, USA
[3]Department of Psychology, University of Oregon, Eugene, OR 97403, USA

*Correspondence:
bkuhl@uoregon.edu (B.A. Kuhl).

Box 1. Methods and Approaches for Detecting Reactivation

Noninvasive Neuroimaging Methods

Memory reactivation has been extensively studied using noninvasive imaging techniques [e.g., fMRI, electroencephalography (EEG), magnetoencephalography], which can simultaneously record population neural activity from large parts of the brain. Below, we detail several popular analytic approaches for measuring reactivation in neuroimaging data. All of these approaches test for similarity in profiles of neural activation during perceptual experience and memory retrieval. Notably, these methods can be applied at varying levels of specificity: stimulus category (e.g., faces vs scenes vs objects), specific exemplar (e.g., apple vs butterfly), or stimulus feature value (e.g., color: red vs yellow).

Univariate Activation

Brain regions are identified for which the mean level of neural activity varies according to the stimulus content. Reactivation is demonstrated by brain regions exhibiting similar content sensitivity (changes in mean activity level) during perception and memory retrieval [19].

Pattern Similarity

Neural activity patterns (activity across voxels or sensors) are measured during perception from a specific brain region. Patterns are then measured within the same brain region during memory retrieval. Reactivation is demonstrated by the correlation between a perception pattern and a corresponding retrieval pattern exceeding the correlation between noncorresponding patterns [2,84].

Decoding Models

Supervised models are trained to learn a mapping between stimuli or stimulus classes and neural activity measured during perception. These models are then tested on neural activity measured during memory retrieval. Reactivation is demonstrated by classifier transfer from perception to retrieval (i.e., the perception-trained model accurately labeling stimuli retrieved from memory) [1,22].

Encoding Models

Models that specify mathematical relationships between physical stimulus properties (e.g., pixel values) and neural activity are fit to neural data measured during perception. Reactivation is demonstrated if model parameters generalize from perception to retrieval (i.e., if the perception parameters can be used to accurately predict brain activity during memory retrieval) [85]. Predicted brain activity can also be used to decode remembered stimuli (encoding–decoding) [24,25] or encoding models can be inverted to reconstruct remembered stimuli (inverted encoding model) [46].

Invasive Physiological and Optogenetic Methods

While the current article focuses on findings from noninvasive neuroimaging methods, which benefit from large-scale spatial coverage of the brain, all of the aforementioned methods can be applied to local field potential or single-cell electrophysiological activity measured invasively in humans (when treated for neurological diseases requiring electrode implants) [86] or in animal experiments. Other approaches, such as large-scale or high-density electrophysiological recordings, are more prevalent in animal models and some approaches are restricted to animal experiments. For example, large single-unit datasets collected from rodents have been essential in characterizing the phenomenon of replay, or the temporally structured reactivation of single-cell firing sequences [62]. Only recently has replay of single-cell firing been shown in humans [87]. Optogenetic techniques, applicable in animal studies solely, have revealed that experimentally induced reactivation in rodents is sufficient to evoke memory behavior [88].

contain more information about the content of an experience when that experience is being retrieved from memory compared with when it is perceptually experienced [14–17]. The clear implication of this is that measures of reactivation, which only test whether activity patterns evoked during retrieval are similar to those evoked during perception, will systematically underestimate the amount, strength, or qualities of content representations during retrieval.

In this opinion article, we argue that there is emerging evidence that content representations during memory retrieval are spatially transformed versions of content representations originally expressed during perception. We use this idea of **spatial transformation** to emphasize that content information: (i) is expressed in different brain regions during perception and memory retrieval and, (ii) that these differences in spatial localization are systematic (or predictable) as opposed to random noise or signal degradation during memory retrieval (Figure 1). While the

Glossary

Content-sensitive patterns of neural activity: patterns of neural activity that reflect ‘what’ is being currently experienced or remembered. Content sensitivity may reflect sensitivity to broad visual categories (e.g., faces vs scenes vs objects), specific exemplars (e.g., apple vs butterfly), or specific feature values of an exemplar (e.g., color: red vs yellow). Content sensitivity can be contrasted with process-level responses that are invariant to the content of an experience.

Default mode network: a network of regions spanning medial and lateral temporal cortex, medial and lateral parietal cortex, and medial prefrontal cortex. These regions show higher activation during internally oriented tasks than externally oriented ones and may play a key role in representing the content of retrieved memories.

Episodic memory: a form of declarative long-term memory that allows for recollection of unique previous experiences.

Memory retrieval: accessing a previously stored event from memory.

Reactivation: phenomenon whereby content-sensitive patterns of neural activity evoked during perception are re-expressed when that experience is remembered. Reactivation has been measured using multiple methodologies (e.g., fMRI, EEG) and analytic approaches (e.g., univariate activation, pattern similarity, decoding models).

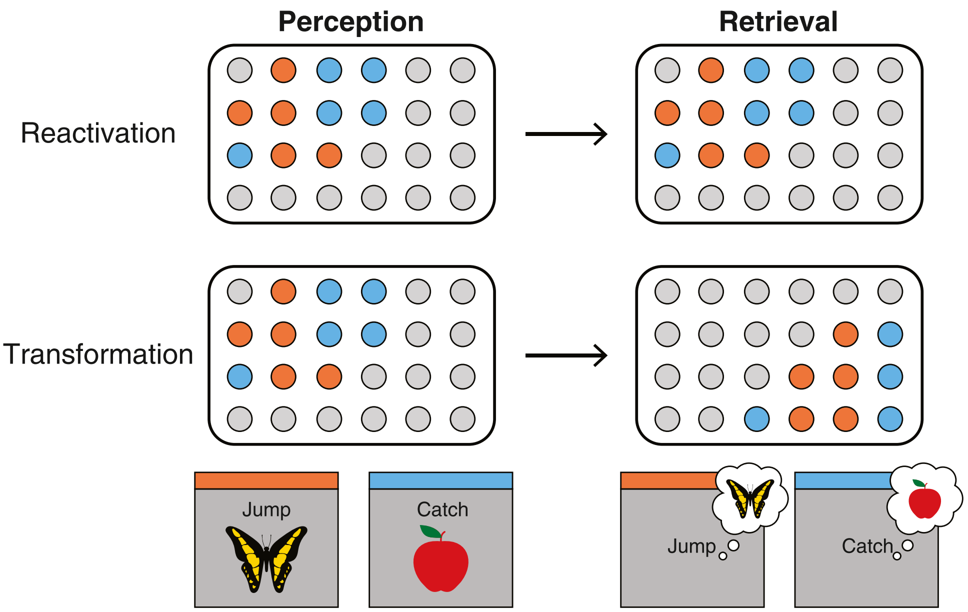
Replay: reactivation of temporally structured sequences of neural activity (most typically, single-cell firing).

Spatial transformation: a form of transformation in which the spatial localization of content-sensitive neural activity patterns changes from perception to memory retrieval. Whereas activity patterns in some brain regions preferentially carry information about perceptual experience, activity patterns in other brain regions preferentially carry information about these same experiences when they are retrieved from memory.

Temporal transformation: a broad class of transformations in which the temporal organization of neural activity changes from perception to memory retrieval.

Transformation: in the context of this article, the term transformation is used to refer to a systematic and predictable change in neural activity from perception to memory retrieval that cannot be attributed to memory noise or error.

Tuning transformation: a form of transformation in which the preferred feature values of an individual neuron, voxel, or population of neurons/voxels changes from perception to memory retrieval.



Trends in Neurosciences

Figure 1. Schematic Comparison of Reactivation and Spatial Transformation. During perception (left column), content-sensitive representations of perceptual events are encoded in cortex. We schematize this by highlighting the units that code for each stimulus (a butterfly and an apple) in different colors. The distinction between reactivation and transformation concerns how these content-sensitive representations at perception relate to content-sensitive representations during memory retrieval (right column). Reactivation (top row) proposes that the representation of an event measured during retrieval will be similar to the representation of the same event measured during perception. We schematize reactivation by highlighting the same neural units active during perception and memory retrieval. Though we illustrate it this way, the idea of reactivation does not argue that there will be a perfect match between representations measured during perception and retrieval; noise or forgetting may degrade this similarity. However, theories of reactivation do not propose or define any systematic differences between perceptual and memory-based representations and typical measures of reactivation do not characterize potential differences. In contrast, the idea of spatial transformation (bottom row) proposes a systematic change in the neural localization of content representations from perception to memory retrieval. We schematize spatial transformation by highlighting different active units during perception and memory retrieval. Critically, spatial transformation cannot be attributed to noise or degradation because the changes are systematic, illustrated here by the rightward shift to a distinct cortical location. Changes in neural localization could occur within cortical regions (e.g., posterior-to-anterior shifts in sensory cortex) or across them (e.g., from sensory cortex to frontoparietal cortex).

possibility of systematic differences in content representations during perception and memory retrieval has, to date, been overshadowed by the more dominant idea of reactivation, it is important to note that cognitive theories of episodic memory have long argued that memory retrieval is more of a constructive act than a reproductive act [18]. Thus, there is a strong theoretical motivation for moving beyond measurements that only index the degree to which content representations at retrieval resemble those at perception and moving toward approaches that characterize how and why content representations at retrieval differ from those at perception. In subsequent sections of this article, we describe some of the recent evidence supporting the specific proposal of spatial transformation from perception to memory. We consider several accounts of why spatial transformations occur and highlight several avenues for future research. We argue that characterizing and understanding perception-to-memory **transformations** (spatial and otherwise) represents a logical, timely, and important next step for the field.

Spatial Transformations within Sensory Cortex

While memory reactivation has been reported in many parts of the human brain, it has been most frequently measured in sensory areas and visual cortex in particular. Early studies demonstrated that visual cortex is active both during visual perception and during retrieval of visual memories [19], demonstrating a coarse-level reactivation of the sensory modality of experience (e.g., vision versus audition). More recently, studies have established reactivation of increasingly fine-grained information within visual cortical areas. For example, numerous studies have shown that patterns of activity in ventral temporal cortex that reflect the visual category of a stimulus (e.g., face versus scene versus object) are reactivated during retrieval of that stimulus [1,20–22]. Further, patterns of neural activity in early visual areas (e.g., V1) that reflect low-level stimulus properties such as spatial location, spatial frequency, edges, and orientation, are also reactivated during retrieval [23–26].

The common thread among these examples of visual cortical reactivation is that the same content-sensitive patterns of activity evoked during perception are re-expressed during memory retrieval. Yet, results from a number of recent studies suggest that there are subtle, but systematic differences in the spatial localization of content information within visual cortex during perception and memory retrieval. In particular, several studies have found that scene-selective responses during perception are located posterior to scene-selective responses during memory retrieval. For example, a meta-analysis of over 100 functional magnetic resonance imaging (fMRI) studies found that scene perception recruits relatively posterior aspects of the parahippocampal place area (PPA), whereas retrieving scenes from memory recruits relatively anterior aspects of PPA [27]. In an unpublished study, this striking dissociation was replicated in a paradigm that directly compared perception and memory retrieval using a within-subjects design [14]. Similar posterior (perception) to anterior (memory retrieval) shifts have been observed in other scene-selective areas, including lateral occipital and medial parietal areas [14,28]. Resting state fMRI analyses have shown that anterior and posterior aspects of ventral, lateral, and medial scene-selective areas have different connectivity profiles with the rest of the brain, consistent with the idea that scene perception and scene memory are supported by segregated visual cortical networks [14,27–29]. Related work has suggested that this posterior-to-anterior division between visual perception and visual memory extends to stimulus classes other than scenes, such as objects, and to aspects of lateral and ventral temporal cortex outside of the scene network [30,31].

Collectively, these studies reveal a shift in the spatial localization of content-sensitive responses during perception and memory retrieval in the visual system. Specifically, information about a stimulus is preferentially expressed in different regions of visual cortex, depending on whether that stimulus is currently available to sensory receptors or recalled from long-term memory. Importantly, the fact that some regions of the visual system contain more information about remembered stimuli than they do about perceived stimuli is fundamentally inconsistent with the idea of reactivation. Moreover, these differences in spatial localization are highly systematic (as opposed to random noise or signal degradation) and generalize across different classes of stimuli and visual regions. As such, these findings strongly suggest that content representations during perception are re-expressed in a spatially transformed manner during memory retrieval.

Spatial Transformations from Sensory Cortex to Frontoparietal Cortex

Though memory reactivation has been most consistently observed in sensory areas, recent applications of pattern-based analysis methods have revealed robust, content-sensitive reactivation in frontoparietal regions [5,7,32,33]. This evidence of frontoparietal reactivation is surprising given the traditional view that frontoparietal regions are involved in controlling or evaluating memories (i.e., memory processes) without actively representing retrieved content [34]. More surprisingly, however, several studies have directly contrasted the strength of content representations during perception and memory retrieval and have found that, among some frontoparietal regions, content representations are not only present during memory retrieval, but are stronger during retrieval than during perception [15–17].

In one study, lateral parietal activity patterns contained more information about the features of visual objects (i.e., the patterns were more correlated within feature than across feature) during memory retrieval than during perception [17]. This contrasted sharply with visual cortical areas, where feature representations were stronger during perception than during memory. Similarly, another study found that representations of individual stimuli were stronger in lateral frontal and lateral parietal regions during memory retrieval than during perception, whereas ventral temporal cortex exhibited the opposite pattern [16]. Strikingly, the fine-grained representational structure among stimuli (i.e., the relative distance between stimuli in representational space) was preserved from perception to memory retrieval despite the change in spatial localization from ventral temporal cortex to frontoparietal cortex. This evidence of preserved information despite changes in spatial localization is strongly consistent with the idea that memory retrieval involves a spatial transformation of content information expressed during perception.

In both of these studies, content representations were stronger in frontoparietal regions during memory retrieval than during perception, a finding that is fundamentally inconsistent with the idea of reactivation. Specifically, ‘perfect’ reactivation should correspond to memory representations that are as strong as, but not stronger than, perceptual representations. Instead, and consistent with posterior/anterior dissociations in visual cortical areas described in the prior section, these findings indicate that some brain regions preferentially represent stimuli or events that are retrieved from memory and that other brain regions preferentially represent stimuli or events that are perceptually available. That said, it is important to note that even in studies that have observed a shift of content representations from visual cortical areas during perception to frontoparietal regions during retrieval, there is still some degree of reactivation both in ventral temporal cortex and in frontoparietal cortex [15–17]. Thus, the presence of transformation does not require the absence of reactivation. Rather, reactivation and transformation likely co-occur. By definition, however, these phenomena explain unique variance in retrieval-related representations. Understanding the factors that determine the relative degree of reactivation versus transformation is a critical open question (see Outstanding Questions).

Why Is There Spatial Transformation?

Why might stimulus representations be spatially transformed from perception to memory? While current evidence does not allow for a definitive answer to this question, we discuss three potential explanations (which are not mutually exclusive) that may help guide future studies.

Distinct Neuroanatomical Origins

One reason why perception- and memory-based representations may be preferentially expressed in different brain regions is because the signals that drive these representations have distinct neuroanatomical origins. While perceptual representations originate from sensory receptors such as those in the retina, retrieved memory representations originate from a pattern completion process commonly thought to be triggered by the hippocampus [35,36]. Although computational perspectives on memory have emphasized the relevance of feedback connections from the hippocampus to sensory cortex in enabling sensory reactivation [37], the connectivity profile of the hippocampus suggests that memory retrieval involves more than sensory reactivation. Most notably, the hippocampus is intimately connected to the **default mode network** [38,39] and this coupling is particularly strong during memory retrieval [40,41]. The default mode network is so-named because regions in this network tend to show higher activation during internally oriented cognition (including memory retrieval) than to perceptual stimuli.

However, several recent studies have found that this bias toward internally oriented cognition also manifests in stronger content representations in the default mode network during memory retrieval than during perception [15–17].

Connectivity with the hippocampus may also explain the anterior bias for memory-based representations in the visual system, as the same anterior regions that exhibit a memory bias also exhibit relatively stronger connectivity to the hippocampus [14,27]. It is also of note that the hippocampus, itself, has been shown to more strongly represent retrieved content than perceived content [16,42]. Thus, just as the hippocampus plays a critical role in driving reactivation in sensory cortex, it is also likely to play a key role in mediating transformation by driving unique cortical responses during memory retrieval that were weak or not present during perception. From this perspective, the degree to which a given cortical area displays a preference for perceptual versus memory-based content may be explained by the degree to which that cortical area is more strongly driven by signals from sensory receptors versus signals from the hippocampus. This idea can be readily tested in targeted studies that combine connectivity measures with measures of content-sensitivity. This account also makes interesting and testable predictions for memory tasks that vary in their dependence on the hippocampus. For example, recognition memory tasks, which only partially depend on hippocampal processing [43], should involve less transformation under this account. Of course, this account leaves open the question of why there are partially distinct neuroanatomical pathways for perception and memory retrieval and whether having segregated content representations is adaptive. Although speculative, it may be the case that segregated representations of perceived and remembered content serve the simple purpose of avoiding confusion between the current environmental state and past environmental states.

Distinct Task Demands

Another potential account of spatial transformation, particularly from visual regions to frontoparietal cortex, is that memory retrieval imposes unique (and perhaps greater) task demands compared with perception. This account is motivated by the known role of frontoparietal regions in implementing attention and control processes [44,45] but specifically õrequires that frontoparietal regions meet task demands by actively representing retrieved content [5,32,46]. For example, frontoparietal regions have been shown to flexibly prioritize retrieved content that is relevant to current memory demands [17,47,48] and to actively insulate retrieved content from sensory distraction [49,50]. From this perspective, whether frontoparietal cortex displays a preference for perception versus memory-based content can be explained by the degree to which content representations must be brought in line with task demands during perception versus memory.

In our view, a ‘task demands’ account does not fully explain existing evidence of transformation. First, this account more readily explains transformations from sensory to frontoparietal cortex than transformations within sensory cortex. Second, this account would predict that the specific frontoparietal regions exhibiting a bias toward memory representations would be the same regions most involved in control or attention. In fact, a bias toward retrieved content is particularly pronounced in regions that belong to the default mode network [16,17], a network that is associated with internally oriented cognition and not top-down attentional control. Finally, task relevance does not appear to be a requirement for frontoparietal regions to exhibit a bias toward memory-based representations. In one study [17], task relevance did influence the strength of memory representations in dorsal regions of parietal cortex (the intraparietal sulcus), but in ventral parietal regions (including components of the default mode network), the strength of memory-based representations was insensitive to task demands, with equivalent representation of mnemonic information that was relevant versus irrelevant to the current task. Strikingly, regions of the rodent parietal cortex have also been shown to contain a bias toward past experience over current perceptual experience, even when past experiences are entirely irrelevant to current behavioral decisions [51]. Taken together, while some frontoparietal regions may be particularly involved in representing retrieved content in a way that aligns them with task demands, this framework, we would argue, is unlikely to fully account for spatial transformations from perception to memory. That said, the 'task demands' account is also testable in that task demands can be independently manipulated during perception and memory retrieval in order to determine whether, or for which brain regions, content representations become stronger as the demand for attention or control increases.

Shift toward Conceptual Representations

Another account of spatial transformation is that the nature of content representations changes from perception to memory. The most obvious version of this account is that, compared with a perceptual representation of an event, a memory-based representation of the same event will reflect relatively more conceptual information. For example, whereas perceiving an apple typically requires at least some low-level processing of stimulus features (contrast, color, spatial frequency, etc.), retrieving the same apple from memory may produce a representation that omits some of these low-level features (i.e., that is less perceptually specific) and is instead biased toward higher-level, conceptual properties derived from general knowledge ('apples are a healthy snack'). This perspective is motivated by evidence that some of the same frontoparietal [52] and anterior ventral temporal [53,54] regions that demonstrate biases toward memory-based representations are also involved in conceptual processing. From this perspective, the degree to which a given cortical area displays a preference for memory-based content should be explained by the degree to which that cortical area codes for conceptual features. Notably, the idea that memories involve a shift toward conceptual representations makes contact with the idea of consolidation, a point we consider in Box 2. Additionally, it is interesting to note that conceptual representations necessarily involve integrating across distinct experiences (e.g., prior encounters with apples), raising the question of whether this integration over time is a relevant quality. Indeed, regions of the default mode network that exhibit biases toward retrieved memories and that have been implicated in conceptual processing have also been shown to have wider temporal receptive windows than sensory regions [55–57].

While an account that emphasizes a shift toward conceptual representations is appealing in many respects, it remains to be seen whether this account can explain recent evidence suggesting spatial transformation in the rodent brain [51] or emerging evidence of spatial transformation in working memory tasks that rely on simple visual stimuli (Box 3). This account also makes a

Box 2. Transformation versus Consolidation

The idea that memories undergo transformation is not new to the field. For instance, numerous prior empirical and theoretical papers have argued that memories are transformed from hippocampally dependent traces to cortically dependent traces via a consolidation process that operates over long timescales [89,90]. This clearly makes contact with the idea of spatial transformation proposed here. It also makes contact with some of the explanations we offer for transformation, namely, consolidation is thought to enable the extraction of conceptual knowledge from individual episodes [90,91]. However, our proposal is also distinct from the standard view of consolidation. In particular, we argue for a rapid transformation wherein perceptual content is systematically reorganized during immediate recall. In contrast, consolidation can be thought of as more gradual transformations of memory-based content over time. That said, the distinction between these ideas may be blurred in some cases. For example, several recent papers have raised the possibility that repeated retrieval accelerates consolidation processes that normally occur over longer timescales [92–94]. Further, a recent paper demonstrated that just one session of retrieval practice (as opposed to restudying information) is sufficient to shift the spatial localization of content representations during subsequent memory retrieval [95]. Thus, it is possible that repeated retrieval may also be a factor that influences the degree of perception-to-memory representations.

Box 3. Parallels to Working Memory

Many neuroimaging studies have demonstrated that visual cortex contains representations of stimuli held in working memory [96,97] and that the quality of these representations relates to behavioral performance in working memory tasks [98,99]. Often referred to as 'sensory recruitment' rather than reactivation, these findings convincingly demonstrate that visual cortex encodes stimuli during a delay period in a similar format to perception. However, recent work has raised the possibility that, in addition to sensory recruitment, sensory representations may also be transformed during working memory maintenance. In particular, one recent study [50] showed that while early visual areas code perceived and maintained orientations in a shared sensory-like format, parietal cortex codes perceived and maintained orientations in distinct formats. While speculative, it is possible that despite the dissociable mechanisms underlying working memory and long-term memory [100,101], common principles of spatial transformation apply across these domains. Indeed, the debate over the functional relevance of working memory representations in sensory versus frontoparietal regions [102,103] in many ways parallels questions and findings related to the functional significance of long-term memory reactivation in sensory versus frontoparietal regions [5,7,16,17,47,104]. Ultimately, understanding the nature of transformation in working memory and long-term memory, and the extent to which they are similar, will require a coordinated effort between working memory and long-term memory researchers as well as direct experimental comparisons across tasks.

notable testable prediction: if spatial transformations reflect differences in the perceptual versus conceptual content of representations, then the degree of transformation should be modulated by the kind of representation accessed during perception and memory retrieval. For example, if perception and retrieval both require conceptual representations, the resulting content representations should be more similar (i.e., relatively less transformation and relatively more reactivation).

Other Forms of Transformation

Though this opinion article focuses on spatial transformation, transformation is likely to occur along other dimensions as well. In this section, we briefly describe other forms of transformation that do not fit our definition of spatial transformation. Although detailed consideration of these other forms of transformation is beyond the scope of the current article, these other lines of evidence reinforce the critical conceptual point that memory representations differ from perceptual representations in systematic ways, differences that strict measures of reactivation will not capture. Ultimately, theoretical accounts of how memories are expressed in the brain during retrieval should integrate all of these lines of work coherently.

First, recent evidence suggests that even when the same neurons or voxels are active during perception and memory retrieval, their tuning for stimulus features may change (**tuning transformation**). For example, the same visual cortical voxels are tuned to more foveal eccentricities and lower spatial frequencies during memory retrieval than perception [58]. These voxels also pool over larger extents of visual space during memory retrieval than perception [58,59]. Similarly, at least some neurons in monkey lateral prefrontal cortex exhibit different motion direction tuning during perception and during working memory maintenance [60]. Other work suggests that electrophysiological responses in sensorimotor and premotor cortex may switch from reflecting sensory tuning during speech perception to reflecting motor tuning during speech rehearsal [61].

Second, a large body of evidence from animal studies, and more recently from human studies, indicates that neural activity undergoes several kinds of **temporal transformations** from perception to memory retrieval. In rodents, it is well established that **replay** of hippocampal place cell sequences involves temporal compression, in that sequences of cell firing occur much faster offline (during rest or sleep) than online (during exploration) [62]. Behaviorally, humans also express memory compression [63,64], and initial evidence for compressed replay in humans has recently been reported using noninvasive measurements of population activity [65–68]. Additionally, memories are influenced by event segmentation, a process that chunks continuous perceptual experience into discrete events [69]. While some segmentation processes influence online perception as well as memory, other segmentations processes occur after perception and may specifically influence how temporal sequences are remembered [70,71]. The extent to which these processes result in activity patterns during retrieval that diverge from earlier perceptual activity is an interesting question for future research. Finally, memory retrieval may involve a reversal of the temporal activity flow present during perception, such that higher-level brain regions engaged during relatively late stages of perceptual processing are invoked during initial stages of memory retrieval [72–74]. These findings and other observations about the temporal structure of reactivation have been recently reviewed in detail [75].

Concluding Remarks

It has long been appreciated that perception and memory retrieval require distinct cognitive processes [76,77] and engage distinct neural structures [78–80]. Yet, measures of memory reactivation fundamentally focus on the similarity of content representations across perception and memory retrieval. While reactivation has been a crucially important phenomenon in the study of memory, we argue that there is a pressing need to better understand the differences between neural representations of perceived and remembered events. Rather than viewing these differences as the product of noise or imprecision in memory retrieval, we argue that there are systematic spatial transformations from perception to memory that are better understood as reformatting of information, either within brain regions or across them. Many open questions remain concerning the nature, causes, and consequences of transformations from perception to memory. These questions represent exciting areas for future research that will hopefully lead to a more comprehensive understanding of how the brain recreates experiences from the past. To conclude, we highlight some specific goals that could guide future research.

One specific target for future studies is to mathematically model the transformation from perceptual to memory-based neural representations. In the studies reviewed here, we have focused on predictable shifts in brain regions that represent perceptual and memory-based content. However, no studies to date (to our knowledge) have specified transfer functions that predict the spatial activity pattern that an individual memory will evoke based on the corresponding activity pattern evoked during perception. Techniques for predicting neural activity patterns across different brain regions have recently been described [81,82] and these techniques can be readily applied to predict perception-to-memory transformations as well. A critical advantage of approaches that yield concrete (mathematical) predictions of spatial activity patterns is that they allow for precisely quantifying the amount of variance in memory-based activity patterns that is explained by perceptual activity patterns in different brain regions (transformation) versus the re-expression of the same perceptual activity patterns (reactivation). Model-based approaches to identifying transfer functions would represent a significant advance over current approaches (reviewed in this article). Model-based approaches could also be extended to quantify the degree to which variables such as task demands and stimulus properties influence the relative degree of transformation versus reactivation (see Outstanding Questions). Finally, comparing transfer functions across individuals will also allow for deeper understanding of the extent to which transformations are idiosyncratic versus shared across individuals [33].

A second important target for future studies is to incorporate finer-grained measures of content representations when considering spatial transformations. The studies reviewed here relied on category-level representations [15,28], exemplar-level representations [16,33] and, to a lesser extent, feature-level representations [17]. The fact that evidence for spatial transformation can be observed at many levels of representational specificity is important, but feature-level representations hold unique appeal in that they allow events to be decomposed such that potentially subtle changes in content representations can be measured. For example, memory-based representations may differ from perceptual representations in terms of dimensionality (number

Outstanding Questions

To what extent do changes in information content account for spatial transformation from perception to retrieval? Are certain stimulus features that are present during perception systematically lost or distorted in memory? Do memory representations gain information that is absent or weakly present during perception through integration with other memories or existing knowledge structures (schemas)?

What determines the relative degree of neural reactivation versus transformation across brain regions observed during memory retrieval? For example, is greater transformation observed when memory tasks promote conceptual processing at retrieval? Conversely, is relatively greater reactivation in sensory areas observed when memory tasks promote perceptual processing? Does the relative degree of reactivation versus transformation depend on whether memory tasks involve recall versus recognition judgments? Do reactivation and transformation trade-off or are they independent?

Does the degree of transformation across brain regions depend on the temporal lag between perception and memory retrieval? Transformation potentially occurs in working memory paradigms with delays on the order of seconds, yet there is also considerable work documenting consolidation-related transformations at timescales of hours to years. What are the similarities and differences between transformations that occur across these vastly different timescales?

What is the relationship between transformation within sensory areas and transformation from sensory to frontoparietal regions? These two forms of transformation have been studied separately to date and it is thus unclear whether they are related and, if so, how. Notably, the frontoparietal and sensory regions that exhibit biases toward memory-based representations are functionally connected with the hippocampus. To what extent can connectivity with the hippocampus explain both sets of findings?

of features) or the weighting of specific features (e.g., perceptual versus conceptual features). Thus, feature-based approaches [17,25] may provide key insight into spatial transformations. Indeed, a complete understanding of how perceptual experiences are transformed into memories will require understanding not only how and when spatial patterns of neural activity differ between perception and memory but also understanding subtle differences in the information contained within those patterns of neural activity.

Finally, while there is some evidence from rodents consistent with our proposal of spatial transformation [51], the ideas described here are largely motivated by human neuroimaging studies. Given that memory reactivation has been extensively studied in both animal models and humans, it will be valuable to test the predictions of spatial transformation more systematically across species. In fact, the ideas we propose here are particularly well suited to cross-species comparisons (so long as multisite recordings are feasible in the animal model) considering that there are well-developed analytic approaches for measuring content representations that transcend recording methodologies and species (e.g., [83]).

Acknowledgments

We thank Zhifang Ye and Maxwell Drascher for helpful discussions and feedback on earlier drafts of this manuscript. S.E.F. was supported by National Institutes of Health (NIH) grant K00-EY031607. B.A.K. was supported by NIH grant R01-NS107727 and National Science Foundation (NSF) CAREER award BCS-1752921.

References

1. Polyn, S.M. *et al.* (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310, 1963–1966
2. Ritchey, M. *et al.* (2013) Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cereb. Cortex* 23, 2818–2828
3. Gordon, A.M. *et al.* (2014) Cortical reinstatement mediates the relationship between content-specific encoding activity and subsequent recollection decisions. *Cereb. Cortex* 24, 3350–3364
4. St-Laurent, M. *et al.* (2015) Distributed patterns of reactivation predict vividness of recollection. *J. Cogn. Neurosci.* 27, 2000–2018
5. Kuhl, B.A. and Chun, M.M. (2014) Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *J. Neurosci.* 34, 8051–8060
6. Kuhl, B.A. *et al.* (2012) Neural reactivation reveals mechanisms for updating memory. *J. Neurosci.* 32, 3453–3461
7. Lee, H. *et al.* (2018) Decomposing parietal memory reactivation to predict consequences of remembering. *Cereb. Cortex* 29, 3305–3318
8. Chanales, A.J.H. *et al.* (2019) Interference between overlapping memories is predicted by neural states during learning. *Nat. Commun.* 10, 5363
9. Tambini, A. and Davachi, L. (2013) Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proc. Natl. Acad. Sci. U. S. A.* 110, 19591–19596
10. Schlichting, M.L. and Preston, A.R. (2014) Memory reactivation during rest supports upcoming learning of related content. *Proc. Natl. Acad. Sci. U. S. A.* 111, 15845–15850
11. Schapiro, A.C. *et al.* (2018) Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nat. Commun.* 9, 3920
12. Rudoy, J.D. *et al.* (2009) Strengthening individual memories by reactivating them during sleep. *Science* 326, 1079
13. Antony, J.W. *et al.* (2012) Cued memory reactivation during sleep influences skill learning. *Nat. Neurosci.* 15, 1114–1116
14. Steel, A. *et al.* (2020) An interface between spatial memory and scene perception in posterior cerebral cortex. *bioRxiv* Published online May 26, 2020. https://doi.org/10.1101/2020.05.25.115147
15. Long, N.M. *et al.* (2016) Hippocampal mismatch signals are modulated by the strength of neural predictions and their similarity to outcomes. *J. Neurosci.* 36, 12677–12687
16. Xiao, X. *et al.* (2017) Transformed neural pattern reinstatement during episodic memory retrieval. *J. Neurosci.* 37, 2986–2998
17. Favila, S.E. *et al.* (2018) Parietal representations of stimulus features are amplified during memory retrieval and flexibly aligned with top-down goals. *J. Neurosci.* 38, 7809–7821
18. Schacter, D.L. *et al.* (1998) The cognitive neuroscience of constructive memory. *Annu. Rev. Psychol.* 49, 289–318
19. Wheeler, M.E. *et al.* (2000) Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11125–11129
20. Ranganath, C. *et al.* (2004) Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J. Neurosci.* 24, 3917–3925
21. O'Craven, K.M. and Kanwisher, N. (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023
22. Kuhl, B.A. *et al.* (2011) Fidelity of neural reactivation reveals competition between memories. *Proc. Natl. Acad. Sci. U. S. A.* 108, 5903–5908
23. Kosslyn, S.M. *et al.* (1995) Topographical representations of mental images in primary visual cortex. *Nature* 378, 496–498
24. Naselaris, T. *et al.* (2015) A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage* 105, 215–228
25. Bone, M.B. *et al.* (2020) Feature-specific neural reactivation during episodic memory. *Nat. Commun.* 11, 1945
26. Bosch, S.E. *et al.* (2014) Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *J. Neurosci.* 34, 7493–7500
27. Baldassano, C. *et al.* (2016) Two distinct scene-processing networks connecting vision and memory. *eNeuro* 3 ENEURO.0178-16.2016
28. Silson, E.H. *et al.* (2019) A posterior–anterior distinction between scene perception and scene construction in human medial parietal cortex. *J. Neurosci.* 39, 705–717
29. Baldassano, C. *et al.* (2013) Differential connectivity within the parahippocampal place area. *NeuroImage* 75, 228–237
30. Lee, S.H. *et al.* (2012) Disentangling visual imagery and perception of real-world objects. *NeuroImage* 59, 4064–4073

31. Bainbridge, W.A. *et al.* (2020) Distinct representational structure and localization for visual encoding and recall during visual imagery. *bioRxiv* Published online April 21, 2020. https://doi.org/10.1101/842120
32. Lee, H. and Kuhl, B.A. (2016) Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. *J. Neurosci.* 36, 6069–6082
33. Chen, J. *et al.* (2016) Shared memories reveal shared structure in neural activity across individuals. *Nat. Neurosci.* 20, 115–125
34. Cabeza, R. *et al.* (2008) The parietal cortex and episodic memory: an attentional account. *Nat. Rev. Neurosci.* 9, 613–625
35. Marr, D. (1971) Simple memory: a theory for archicortex. *Phil. Trans. R. Soc. B Biol. Sci.* 262, 23–81
36. O'Reilly, R.C. and McClelland, J.L. (1994) Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682
37. Treves, A. and Rolls, E.T. (1994) Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4, 374–391
38. Kahn, I. *et al.* (2008) Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *J. Neurophysiol.* 100, 129–139
39. Ritchey, M. and Cooper, R.A. (2020) Deconstructing the posterior medial episodic network. *Trends Cogn. Sci.* 24, 451–465
40. Huijbers, W. *et al.* (2011) The hippocampus is coupled with the default network during memory retrieval but not during memory encoding. *PLoS One* 6, e17463
41. Higgins, C. *et al.* (2020) Replay bursts coincide with activation of the default mode and parietal alpha network. *bioRxiv* Published online June 24, 2020. https://doi.org/10.1101/2020.06.23.166645
42. Lee, S.-H. *et al.* (2019) Differential representations of perceived and retrieved visual information in hippocampus and cortex. *Cereb. Cortex* 29, 4452–4461
43. Eichenbaum, H. *et al.* (2007) The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* 30, 123–152
44. Miller, E.K. (2000) The prefrontal cortex and cognitive control. *Nat. Rev. Neurosci.* 1, 59–65
45. Colby, C.L. and Goldberg, M.E. (1999) Space and attention in parietal cortex. *Annu. Rev. Neurosci.* 22, 319–349
46. Ester, E.F. *et al.* (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* 87, 893–905
47. Kuhl, B.A. *et al.* (2013) Dissociable neural mechanisms for goal-directed versus incidental memory reactivation. *J. Neurosci.* 33, 16099–16109
48. Long, N.M. and Kuhl, B.A. (2018) Bottom-up and top-down factors differentially influence stimulus representations across large-scale attentional networks. *J. Neurosci.* 38 2724–17
49. Bettencourt, K.C. and Xu, Y. (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat. Neurosci.* 19, 150–157
50. Rademaker, R.L. *et al.* (2019) Coexisting representations of sensory and mnemonic information in human visual cortex. *Nat. Neurosci.* 22, 1336–1344
51. Akrami, A. *et al.* (2018) Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* 554, 368–372
52. Jeong, S.K. and Xu, Y. (2016) Behaviorally relevant abstract object identity representation in the human parietal cortex. *J. Neurosci.* 36, 1607–1619
53. Peelen, M.V. and Caramazza, A. (2012) Conceptual object representations in human anterior temporal cortex. *J. Neurosci.* 32, 15728–15736
54. Martin, C.B. *et al.* (2018) Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *eLife* 7, e31873
55. Hasson, U. *et al.* (2015) Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn. Sci.* 19, 304–313
56. Lerner, Y. *et al.* (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* 31, 2906–2915
57. Baldassano, C. *et al.* (2017) Discovering event structure in continuous narrative perception and memory. *Neuron* 95, 709–721
58. Breedlove, J.L. *et al.* (2020) Generative feedback explains distinct brain activity codes for seen and mental images. *Curr. Biol.* 30, 2211–2224
59. Favila, S.E. *et al.* (2020) Perception and memory have distinct spatial tuning properties in human visual cortex. *bioRxiv* Published online April 30, 2020. https://doi.org/10.1101/811331
60. Mendoza-Halliday, D. and Martinez-Trujillo, J.C. (2017) Neuronal population coding of perceived and memorized visual features in the lateral prefrontal cortex. *Nat. Commun.* 8, 1–13
61. Müsch, K. *et al.* (2020) Transformation of speech sequences in human sensorimotor circuits. *Proc. Natl Acad. Sci. USA* 117, 3203–3213
62. Ólafsdóttir, H.F. *et al.* (2018) The role of hippocampal replay in memory and planning. *Curr. Biol.* 28, R37–R50
63. Arnold, A.E.G.F. *et al.* (2016) Mental simulation of routes during navigation involves adaptive temporal compression. *Cognition* 157, 14–23
64. Bonasia, K. *et al.* (2016) Memory and navigation: compression of space varies with route length and turns. *Hippocampus* 26, 9–12
65. Liu, Y. *et al.* (2019) Human replay spontaneously reorganizes experience. *Cell* 178, 640–652
66. Michelmann, S. *et al.* (2019) Speed of time-compressed forward replay flexibly changes in human episodic memory. *Nat. Hum. Behav.* 3, 143–154
67. Schuck, N.W. and Niv, Y. (2019) Sequential replay of nonspatial task states in the human hippocampus. *Science* 364, eaaw5181
68. Wimmer, G.E. *et al.* (2020) Episodic memory retrieval success is associated with rapid replay of episode content. *Nat. Neurosci.* 23, 1025–1033
69. Kurby, C.A. and Zacks, J.M. (2008) Segmentation in the perception and memory of events. *Trends Cogn. Sci.* 12, 72–79
70. Clewett, D. *et al.* (2019) Transcending time in the brain: How event memories are constructed from experience. *Hippocampus* 29, 162–183
71. Sols, I. *et al.* (2017) Event boundaries trigger rapid memory reinstatement of the prior events to promote their representation in long-term memory. *Curr. Biol.* 27, 3499–3504
72. Linde-Domingo, J. *et al.* (2019) Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat. Commun.* 10, 1–13
73. Griffiths, B.J. *et al.* (2019) Directional coupling of slow and fast hippocampal gamma with neocortical alpha/beta oscillations in human episodic memory. *Proc. Natl Acad. Sci. USA* 116, 21834–21842
74. Staresina, B.P. *et al.* (2019) Recollection in the human hippocampal-entorhinal cell circuitry. *Nat. Commun.* 10, 1503
75. Staresina, B.P. and Wimber, M. (2019) A neural chronometry of memory recall. *Trends Cogn. Sci.* 23, 1071–1085
76. Craik, F.I.M. *et al.* (1996) The effects of divided attention on encoding and retrieval processes in human memory. *J. Exp. Psychol. Gen.* 125, 159–180
77. Fernandes, M.A. and Moscovitch, M. (2000) Divided attention and memory: evidence of substantial interference effects at retrieval and encoding. *J. Exp. Psychol. Gen.* 129, 155–176
78. McDermott, K.B. *et al.* (1999) Direct comparison of episodic encoding and retrieval of words: an event-related fMRI study. *Memory* 7, 661–680
79. Prince, S.E. *et al.* (2005) Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations. *J. Neurosci.* 25, 1203–1210
80. Spaniol, J. *et al.* (2009) Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. *Neuropsychologia* 47, 1765–1779
81. Cole, M.W. *et al.* (2016) Activity flow over resting-state networks shapes cognitive task activations. *Nat. Neurosci.* 19, 1718–1726
82. Ward, E.J. *et al.* (2018) General transformations of object representations in human visual cortex. *J. Neurosci.* 38, 2800–2817
83. Kriegeskorte, N. *et al.* (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60, 1126–1141

84. Staresina, B.P. *et al.* (2012) Episodic reinstatement in the medial temporal lobe. *J. Neurosci.* 32, 18150–18156
85. Senden, M. *et al.* (2019) Reconstructing imagined letters from early visual cortex reveals tight topographic correspondence between visual mental imagery and perception. *Brain Struct. Funct.* 224, 1167–1183
86. Schreiner, T. and Staudigl, T. (2020) Electrophysiological signatures of memory reactivation in humans. *Phil. Trans. R. Soc. B Biol. Sci.* 375, 20190293
87. Vaz, A.P. *et al.* (2020) Replay of cortical spiking sequences during human memory retrieval. *Science* 367, 1131–1134
88. Josselyn, S.A. and Tonegawa, S. (2020) Memory engrams: recalling the past and imagining the future. *Science* 367, eaaw4325
89. Alvarez, P. and Squire, L.R. (1994) Memory consolidation and the medial temporal lobe: a simple network model. *Proc. Natl. Acad. Sci. U. S. A.* 91, 7041–7045
90. Nadel, L. *et al.* (2000) Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* 10, 352–368
91. McClelland, J.L. *et al.* (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457
92. Brodt, S. *et al.* (2016) Rapid and independent memory formation in the parietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 113, 13251–13256
93. Brodt, S. *et al.* (2018) Fast track to the neocortex: a memory engram in the posterior parietal cortex. *Science* 362, 1045–1048
94. Antony, J.W. *et al.* (2017) Retrieval as a fast route to memory consolidation. *Trends Cogn. Sci.* 21, 573–576
95. Ye, Z. *et al.* (2020) Retrieval practice facilitates memory updating by enhancing and differentiating medial prefrontal cortex representations. *eLife* 9, e57023
96. Harrison, S.A. and Tong, F. (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635
97. Serences, J.T. *et al.* (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* 20, 207–214
98. Emrich, S.M. *et al.* (2013) Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *J. Neurosci.* 33, 6516–6523
99. Sprague, T.C. *et al.* (2014) Reconstructions of Information in visual spatial working memory degrade with memory load. *Curr. Biol.* 24, 2174–2180
100. Shallice, T. and Warrington, E.K. (1970) Independent functioning of verbal memory stores: a neuropsychological study. *Q. J. Exp. Psychol.* 22, 261–273
101. Alvarez, P. *et al.* (1994) The animal model of human amnesia: long-term memory impaired and short-term memory intact. *Proc. Natl. Acad. Sci. U. S. A.* 91, 5637–5641
102. Xu, Y. (2018) Sensory cortex is nonessential in working memory storage. *Trends Cogn. Sci.* 22, 192–193
103. Scimeca, J.M. *et al.* (2018) Reaffirming the sensory recruitment account of working memory. *Trends Cogn. Sci.* 22, 190–192
104. Jonker, T.R. *et al.* (2018) Neural reactivation in parietal cortex enhances memory for episodically linked information. *Proc. Natl. Acad. Sci.* 115, 11084–11089