The cognitive neuroscience of memory representations

Michael D. Rugg [a], Louis Renoult [b],[*] (ID)

[a] *Center for Vital Longevity and School of Behavioral and Brain Sciences, The University of Texas at Dallas, Dallas, TX 75235, USA*
[b] *School of Psychology, University of East Anglia, Norwich, Norfolk, United Kingdom*

ABSTRACT

The present paper considers the cognitive neuroscience of memory from a representational perspective with the aim of shedding light on current empirical and theoretical issues. We focus on episodic memory, differentiating active versus latent, and cognitive versus neural memory representations. We adopt a causal perspective, according to which a memory representation must have a causal connection to a past event to count as a memory. We note that retrieved episodic information may nonetheless only partially determine the content of an active memory representation, which can comprise a combination of the retrieved information with semantic, schematic and situational information. We further note that, especially in the case of memories for temporally remote events, re-encoding operations likely lead to a causal chain that extends from the original experience of the event to its currently accessible memory trace. We discuss how the reinstatement framework provides a mechanistic basis for the causal linkage between an experience, the memory trace encoding it, and the episodic memory of the experience, highlighting the crucial role of hippocampal engrams in encoding patterns of neocortical activity that, when active, constitute the neural representation of an episodic memory. Finally, we discuss some of the ways in which a memory can become modified and hence distanced from the episode that precipitated it.

1. Introduction

In this paper we discuss the concept of a memory representation in relation to the cognitive neuroscience of long-term memory. We discuss the ways in which cognitive and neural representations and related concepts are used in current cognitive neuroscientific theories of memory, with a focus on episodic memory (consciously accessible memories for personally experienced unique events; Tulving, 1983). We touch on some of the problematic issues that arise from this usage and tentatively suggest adjustments to current terminology that might help to resolve these issues. We consider how notions of memory representations framed in neural and cognitive terms might best be articulated, and how these notions might be expanded to incorporate semantic memory (conceptual and factual knowledge). We conclude by discussing the different factors that might modify or transform a memory representation.

The scope of the paper is relatively narrow: our focus is on what is historically referred to as 'declarative' memory in cognitively healthy adults. Thus, we do not discuss important topics such as memory impairment or development and nor do we discuss non-declarative

(implicit) memory. While we briefly discuss what we think to be relevant ideas drawn from the philosophy of memory, we acknowledge that our acquaintance with this large and technically intimidating literature is limited and make no claims for thoroughness or originality in this respect.

2. Conceptual underpinnings

2.1. Mental, neural and memory representations

The questions of what in the mind constitutes a 'representation', and the role that representations play in cognition, have long exercised philosophers and, more recently, cognitive scientists and psychologists also (Pitt, 2022). Much of this inquiry has focused on the role of mental representations – mental entities that are held to have an intentional[1] relation with something else (Brentano, 1874). Mental representations can not only be about things that are existent (my dog, for example) but also things that have never existed (e.g., phlogiston) or, particularly relevant here, things that no longer exist, such as my birthday party last week. From this perspective, a memory representation has 'content', but

---

[1] The term 'intentional' is employed here in its technical, philosophical, sense. A mental representation is intentional because it 'stands for' or is 'about' something.

the content applies to a non-existent 'object' (see De Brigard, 2023, for discussion of this and related issues).

The concept of a representation has also been applied at the neural level, when representations might be defined as spatiotemporal patterns of neural activity that encode information about some external or internal input and have the potential to influence behavior and, in some cases, awareness (e.g., Baker et al., 2022; Dudai, 2007; Kriegeskorte and Diedrichsen, 2019; Poldrack, 2021). The distinction drawn here between mental and neural representations should not be taken to imply that mental (hereafter, 'cognitive') representations are disconnected from the brain – we take it as axiomatic that all cognitive phenomena are caused by neurobiological processes. As we discuss below in our discussion of memory representations, however, the mapping between a cognitive memory representation and its neural substrate(s) can be complex.

2.2. Active versus latent representations

An ambiguity arises from the common practice in the memory field of using the term 'representation' to mean two different things, which we distinguish here as referencing 'active' as opposed to 'latent' representations (cf. Frankland et al., 2024). Active memory representations refer to information-bearing entities that support the maintenance of a retrieved memory in a state that allows the memory to influence cognition and behavior. In the case of a cognitive memory representation, this usage usually refers to information corresponding to mnemonic content that can be attended, reported on and employed to guide goal-directed behavior[2] – that is, there is something 'it is like' to have an active memory representation. At the neural level, an active representation refers to a neural ensemble, which might be highly distributed, that demonstrates a pattern of activity that correlates[3] at least partially with the pattern that was elicited by, and contributed to the neural representation of a past experience (see 'retrieval-related reinstatement below), and which is necessary for a memory of the experience to influence on-going cognition and behavior. There is no reason to suppose that every active neural memory representation has a corresponding cognitive representation or, at least, a consciously accessible one, as is evidenced by the extensive literature on implicit (non-declarative) memory (Schacter et al., 1993)[4] and by the finding that memory representations can be activated during sleep (Hu et al., 2020). By contrast, what we call a latent representation corresponds to what might be described as a 'dormant' or 'potential' memory, sometimes referred to as a memory 'trace' or, in the terminology of Semon (1904), an 'engram'. It is unclear to us that the notion of a latent representation can be made coherent at the cognitive level; rather, it seems to be meaningful at the neural level only (Norman, 2010). By this view, a latent representation exists when an event elicits a change in the brain that encodes information about an experience in a manner that allows at least some of the information to later be (re)activated and to contribute to the content of an active memory representation.

The distinction between active and latent representations raises the question of how a representation is 'converted' or 'transformed' from a latent to an active state. In the case of episodic memory, it is frequently proposed that memory retrieval depends on a pattern completion process, which occurs when a retrieval cue triggers the reactivation of the neural population that encoded the episode when it was initially experienced (see Section 3.1). Whether this process is invariably 'all or none' (cf. Norman, 2010), or whether instead it can be graded, leading to episodic memory representations that vary in 'strength' (Ingram et al., 2012), is an unresolved question.

Whereas the distinction between active and latent representations might seem obvious, we think there is potential for confusion when a single word – representation – is employed to refer to both concepts. This potential is not merely hypothetical: for example, one of the present authors managed to employ each usage multiple times in the course of a single sentence, '….*pre-retrieval processes responsible for generating a representation of a retrieval cue that maximizes overlap with a stored memory representation; retrieval ('ecphoric') processes that occur when a cue representation overlaps sufficiently with a stored (latent) memory representation to cause 'pattern completion' and the reinstatement of the representation in the cortex*…' (Rugg, 2024, p. 1062). This unfortunate conflation of meanings is somewhat mitigated by the qualifiers 'stored' and 'latent' when referencing non-active representations but, even so, hindsight suggests that more consistent terminology would have added clarity to these statements. Accordingly, in the rest of the current paper, the term 'representation' will refer exclusively to *active* representations (whether cognitive or neural); latent representations will be referred to as 'traces'[5]; or *engrams*, consistent with Semon's original definition of an engram ('….the enduring though primarily latent modifications in the irritable substance produced by a stimulus….'; Semon, 1904, p.12; see also Josselyn and Tonegawa, 2020).

2.3. What counts as a memory?

Although the answer to this question might seem obvious, philosophers would disagree (see, for example, De Brigard, 2023; Michaelian and Sutton, 2017). According to one view, for example, there is no necessity for a cognitive representation to have a causal connection with a specific past event[6] to count as an episodic memory; all that is required is that the representation is interpreted as a memory, that is, it is 'past-oriented' (e.g., Michaelian, 2016). Like, we suspect, many other memory researchers we are unsympathetic to this view, which would seem to imply an equivalence between a veridical memory and a non-veridical memory, irrespective of how much the latter deviates from veracity. Rather, we incline toward the position referred to in the philosophical literature as the 'causal approach' (e.g., Debus, 2017; Martin and Deutscher, 1966; Michaelian and Robins, 2018) and draw a distinction between memory and other kinds of cognitive representations, including those that are 'wrongly' interpreted as memories. According to this position, a necessary, though not a sufficient condition for a cognitive representation to count as a memory is that it is causally

---

[2] We acknowledge that this kind of language invites the accusation of succumbing to the 'homuncular fallacy' – the implication that some unspecified intelligent agent is doing the 'attending', 'reporting', etc. Whether and how this can be avoided is an important question beyond the scope of the present paper (see, e.g., Guest and Martin, 2025).

[3] We use the term 'correlates with' rather 'resembles' because of the possibility that the same or similar content (here, an experience and a later memory of it) could be represented in different neural populations.

[4] It is however questionable whether the neural mechanisms underpinning some kinds of implicit memory (e.g. procedural memory or priming) would qualify as neural representations if such representations are required to be content bearing.

[5] This usage of 'trace' as synonymous with engram is only one of the ways that the term has been interpreted and used in the philosophical and psychological literature; Robins, (2025).

[6] The definition of an 'event' (episode) is murky. Events can range from a temporally circumscribed occurrence such as the appearance of a study item during a laboratory experiment (what Tulving, 1983, referred to as a 'miniature-event') to temporally extended experiences such as viewing a movie or going on vacation. See Zacks (2020) for a recent review of the relationship between event perception and memory.
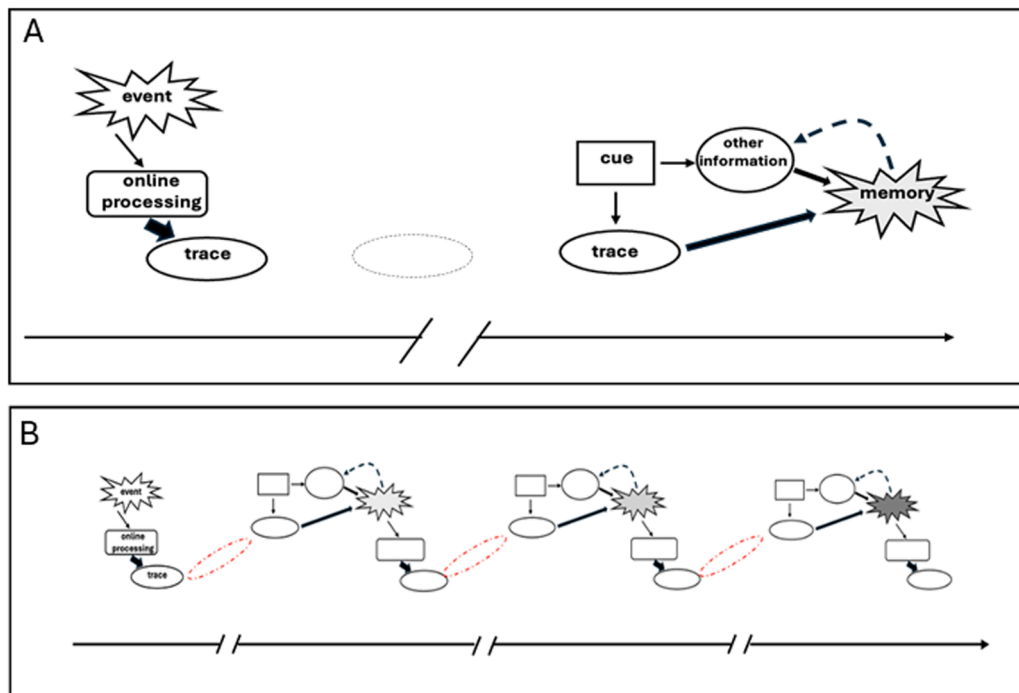
Fig. 1. A. Schematic depiction of the encoding (left) and subsequent retrieval (right) of an episodic memory according to the modified causal theory described in the text. A record of the processing accorded an event is encoded in a memory trace which remains dormant (dotted outline) until activated by a retrieval cue. The activated trace reinstates the record of the processing of the event, which forms much of the content of a corresponding memory representation. The representation also includes a variable amount of non-episodic (e.g. schematic) information elicited both by the cue and the retrieved mnemonic content. B. Schematic of a causal chain resulting from iterative re-encoding. The differing hues of the shape symbolizing the retrieved memory representations reflect their gradual modification as a consequence of the multiple retrieval-encoding cycles.

connected to a past event; that is, the representation would not exist if the event had not been experienced,[7],[8] Note that as we interpret it, the causal definition of memory does not require that memory content is determined exclusively by a unique event (hence, our position deviates from what has been called the 'classical' causal approach; Michaelian and Robins, 2018; see also Andonovski, in press). As we discuss later, memory content is often an amalgam of information about an event and other, more generic information. That is, as has been discussed many times previously (e.g., Addis, 2020; Conway, 2009; Irish, 2019; Schacter et al., 2007), the information encoded about an event need not be the sole determinant of the content of a later retrieved memory of it (Fig. 1A). Furthermore, as we also elaborate below, our interpretation of the causal definition of a memory representation allows for the possibility that memory for an event, and its corresponding memory trace, might depend on a causal *chain*. This interpretation acknowledges the widely held notion that a retrieved memory can be 're-encoded', preserving much of its content while reinforcing, replacing or even duplicating its neural substrate (the latent neural representation or memory trace; note that the last of these possibilities is what is implied by 'multiple trace theory', Nadel and Moscovitch, 1997; see Section 4.1.3.). Thus, for any memory that has undergone re-encoding, there can be a mismatch between the age of the memory trace (which dates back to the last time the memory was retrieved) and the age of the event that is represented by the trace; that is, crudely speaking, there is a mismatch

between the age of the 'vehicle' of the memory and its 'content' (see, for example, Dretske, 2003 for considerably more sophisticated discussion of this distinction in relation to mental representations). Moreover, especially in the case of memories for temporally remote events, re-encoding can be iterative (Moscovitch and Gilboa, 2024), with each re-encoding operation potentially adding another link to the causal chain that extends from the experienced event to its currently accessible memory trace (Fig. 1B).

A causal definition of memory is implicit in innumerable laboratory-based studies of episodic memory in which participants are exposed to a list of study items and subsequently tested for their ability to recall or recognize the items. In such studies, only responses that correctly identify items as having been studied are taken as evidence of successful memory retrieval; regardless of any accompanying phenomenology 'false recalls' or 'false alarms' (items incorrectly endorsed as previously studied) are classified as memory errors. The causal definition is also applicable to more complex situations such as retrieval of source (contextual) information or, more complex still, temporally remote memories, although in this latter case verification of a causal link between a reported memory and the event it purports to correspond with can be challenging. According to the position adopted here, in the absence of such a causal link, a reported memory should be classified as a case where the 'recalled' information was *imagined* rather than *remembered*. This distinction raises the question of what in the brain distinguishes between the retrieval of a veridical as opposed to a non-veridical (imagined) memory, especially since the underlying cognitive representations seemingly can be phenomenologically indistinguishable (Bernstein and Loftus, 2009). As we elaborate below, a possible answer to this question is that veridical memories are the manifestations of active representations generated by (re)activation of the memory trace (engram) that was formed when an event was

---

[7] The concept of causality raises profound issues that have exercised philosophers for centuries, and continue to do so (e.g. Gallow, 2022). For present purposes, however, we adopt the rather simplistic counterfactual definition of a cause implied in the text: i.e. *X* is a cause of *Y* if *Y* would not have occurred had *X* not occurred (e.g., Lewis, 1973; Menzies and Beebee, 2025).

[8] Another condition that arguably must be satisfied is that the representation is attributed to the past and not to current cognitive processing; cf. 'unconscious plagiarism'; Taylor, (1965).
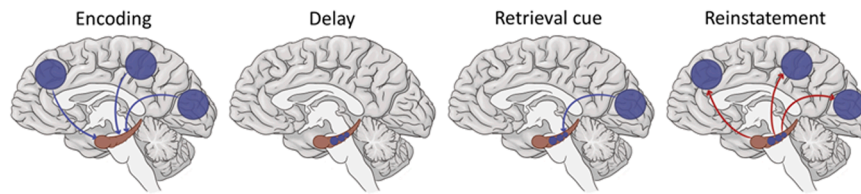
Fig. 2. Episodic memory retrieval involves the reactivation of the cognitive and neural processes which were active when the event was initially experienced. At encoding, neural patterns are indexed and stored by the hippocampus, and hippocampally mediated pattern separation processes ensure that similar episodic events are stored as non-overlapping memory representations. Later, a retrieval cue which partially overlaps with the neural pattern elicited by the original event triggers hippocampal pattern completion, which gives rise to the reinstatement (reactivation) of the neural activity originally elicited by the episode at encoding (Reproduced with the authors' permission from Rugg and Srokova, 2024).

experienced or, perhaps, last retrieved.[9] A potential obstacle to this line of argument is the existence non-veridical memory judgments that arise from memory misattributions – for example, prior list intrusions during free recall. In our view, however, these sorts of memory errors are better construed as a failure of post-retrieval memory *monitoring* rather than faulty memory *retrieval* (Johnson et al., 1993; for review see Rugg, 2024). By this argument, these errors arise when an otherwise veridical memory is attributed to the 'wrong' cause.

2.4. How does the causal definition apply to semantic memory?

The foregoing discussion focused on episodic memory but can be extended to the consideration of semantic memory. Semantic memories are widely held to arise from the distillation of statistical regularities across multiple episodes that contain overlapping information (Renoult et al., 2019), although it has been argued that they can also originate from a single episode (e.g., Gilboa and Moscovitch, 2021; McClelland et al., 1995). Therefore, semantic memories are causally linked to the past no less than episodic memories. Unlike in the case of episodic memory, however, the causal link ultimately may not be with a specific prior episode but with multiple episodes (thereby violating a condition of the 'classical' causal approach – that of a one-to one mapping between an event and the memory of it, Michaelian and Robins, 2018). As has been suggested previously (e.g., Kumar, 2021), this aspect of semantic memories is likely responsible for their acontextual quality and the absence of a phenomenological sense of remembering when they are retrieved. An important and much discussed question, which we return to below, concerns how an episodic memory becomes incorporated into semantic memory.

3. Cognitive neuroscience of memory representations

Below, we build on the foregoing discussion and review research relevant to the goal of developing a cognitive neuroscientific account of memory that is informed by the ideas about neural and cognitive representations discussed above. We begin by describing and extending a long-standing framework that provides a mechanistic basis for the causal linkage between an experience, the memory trace encoding it, and retrieval of an episodic memory of the experience. We also discuss some of the ways in which a memory can become modified and hence distanced from the episode that precipitated it.

3.1. Retrieval-related reinstatement

There is a strong consensus that the content of an episodic memory representation emerges from distributed patterns of neural activity in brain regions that played a role in representing the content during its 'on-line' processing. This general idea receives support from the

extensive literature describing studies of memory 'reinstatement' in humans with functional magnetic resonance imaging (fMRI) and related non-invasive neuroimaging methods (for reviews, see Danker and Anderson, 2010; Favila et al., 2020; Rugg et al., 2008). Of importance, reinstatement has been demonstrated not only in studies employing individual items such as words and images, but also when more 'naturalistic', multi-modal materials such as movies are employed (Chen et al., 2017; Oedekoven et al., 2017; Reagh and Ranganath, 2023).

Reinstatement (synonymously, reactivation) refers to the idea that memory retrieval depends on the re-engagement of neural processes that were active when the memory was first encoded. The idea has a long history. An early version was proposed by William James (James, 1914) and it was articulated in a remarkably modern form by Semon (1904). Reinstatement is central to contemporary research on 'engram cells' (see below) and, more relevant to current concerns, it is an important component of several historically influential models of hippocampally-mediated memory retrieval (e.g., Alvarez and Squire, 1994; Marr, 1971; Teyler and Discenna, 1985). In these and related models, memory retrieval (what Semon referred to as 'ecphory') occurs when the pattern of neural activity elicited by an event when it was experienced is reinstated by the activation of a hippocampally stored engram encoding the pattern. Through this mechanism, the neural populations that were concurrently active during the on-line processing of the event are co-activated during its retrieval, reinstating associations between features of the event that, together, make it distinct from other similar occurrences. Because the hippocampus is highly adept at 'pattern separation' and 'pattern completion' (Yassa and Stark, 2011), retrieval cues eliciting neural activity that only partially overlaps with the activity elicited by the original event can be sufficient to elicit activation of the entire neural ensemble representing the pattern of cortical activity elicited as the event was encoded and hence to reinstatement of the activity. By this view, therefore, hippocampal engrams do not directly encode memories of prior experiences. Rather, they encode the patterns of neocortical activity that, when active, constitute the neural representations of those memories. Thus, hippocampal engrams can be thought of as 'second order' representations (representations of representations). Consequently, an active hippocampal memory representation does not have 'content' in the conventional sense but instead provides the 'instructions' that allow the neocortex to actively represent mnemonic content.[10] This perspective provides a neural basis for an important psychological principle embraced by memory researchers belonging to the 'proceduralist' tradition (e.g., Kolers, 1973; Morris et al., 1977): what gets encoded into memory is not an 'objective' depiction of an event, but rather a record of the processing that the event was accorded. That is, memories are a 'by-product' of the on-line

---

[9] As noted previously, the memory trace might have been formed when the event was last retrieved, rather than concurrently with the original experience.

[10] The focus here on the hippocampus should not be taken to imply that the structure acts alone during memory retrieval. There is evidence that the hippocampus is a key member of an extended brain network that includes multiple neocortical and subcortical regions, the collective function of which is to support successful episodic memory retrieval (see, e.g., Rugg and Vilberg, 2013; Ritchey et al., 2015).

processing that is engaged as we interact with the world (Craik and Lockhart, 1972; see also Rugg et al., 2008; Rugg et al., 2015). Fig. 2 illustrates one way these ideas can be schematized in terms of large-scale patterns of brain activity.

To the extent that episodic retrieval invariably involves reinstatement of neural activity that was elicited when the episode was experienced, reinstatement provides a means (at least in principle) of distinguishing between 'memories' corresponding to prior rather than imagined events, regardless of the accompanying phenomenology: reinstatement should only be evident in the former case, a consequence of the activation of the memory trace (engram) that encoded the pattern of activity associated with the event now being remembered. Of course, as a memory becomes distanced from the event that was its original cause through one or more of the mechanisms discussed in Section 4, the completeness of reinstatement will likely decline, ultimately perhaps to a point where the causal link with the original precipitating event is effectively broken. At this point, the event might be forgotten (corresponding to retrieval of no or minimal relevant mnemonic content) or, alternately, a 'memory' representation that comprises content constructed from generic information might be misattributed to the event. Similarly, a memory might become temporally distanced from its precipitating event by virtue of re-encoding (Section 2.3). In this case reinstatement would presumably reflect not the pattern of activity elicited when the episode was initially experienced, but the pattern associated with the most recent prior retrieval. To the extent that re-encoding is associated with modification of the original mnemonic content, a point might be reached where the causal chain originating with the original episode is broken.

For a variety of reasons, Fig. 2 should not be interpreted too literally (see Rugg et al., 2015 for a fuller discussion). In the first place, overlap between encoding- and retrieval-related neural activity must be differentiated on a much more fine-grained scale than the regional level depicted in the figure (likely at the single neuron level). Thus, the task of the hippocampus, or any other structure whose function is to capture and later reinstate patterns of activity associated with the processing of an event, is considerably more complex than merely registering the neocortical areas that were co-activated as the event was processed.

Second, Fig. 2 is far from a complete specification of the component processes supporting successful retrieval: if this is all there was, it is unclear how we would be able to distinguish between the online experience of an event and a later memory of it. Furthermore, if episodic retrieval occurred every time there was overlap between current and past processing, we would be in a state of almost continuous retrieval. As was noted by Tulving (1983), these and related considerations imply that episodic retrieval must be under some kind of control. Tulving proposed that stimulus events are processed as retrieval cues only when the rememberer adopts a specific cognitive state, which he called 'retrieval mode'. According to this proposal, depending on whether retrieval mode is adopted, the same stimulus event can be processed either as an episodic retrieval cue or with respect to its online significance. Additionally, while reinstatement might contribute to the content of a memory representation (see below), the question of how that content is accessed in service of cognition and behavior is not addressed within the framework outlined in Fig. 2 (cf. King et al., 2015; Rugg, 2024; Rugg and Vilberg, 2013). Indeed, recent evidence indicates that the content of a retrieved memory representation is itself subject to control: as indexed by retrieval-related reinstatement effects, young adult participants appear capable of retrieving only those features of an encoded episode that are relevant to the current retrieval goal, (a process that has been referred to as 'retrieval gating'; de Chastelaine et al., 2025; Elward and Rugg, 2015; Srokova et al., 2021).

A further qualification arises from the fact that Fig. 2 implies that retrieval comprises little more than recapitulation of the processing engaged by the original experience. If this were so, then recollection presumably would be 'all or nothing'; either everything that was registered in the brain as an event unfolded would be retrieved, or nothing

would be. Moreover, memories would be largely veridical. Clearly, neither of these scenarios is accurate. Memories can be an imperfect mirror of experience and are frequently partial and abstracted records of the originally experienced event. Among the factors contributing to the incomplete relationship between an event and a later memory of it, three stand out. First, different features of an event are not equally likely to be encoded. Other things equal, features that are attended the most fully are the ones most likely to be later remembered (e.g., Chun and Turk-Browne, 2007; Uncapher and Rugg, 2009). Relatedly, temporally extended episodes appear to be encoded as compressed segments that are separated by 'event boundaries', with an emphasis on information occurring close to the boundaries (Jeunehomme and D'Argembeau, 2019, 2020; Zacks, 2020). Thus, memories are not a continuous record of experience. Lastly, as discussed in more detail in Section 4, episodic retrieval is frequently constructive such that retrieved information is combined with non-episodic information and the outcome interpreted in the context of current expectations and biases (e.g., Bransford and Franks, 1971; Brewer, 1986; Brewer and Treyens, 1981; for reviews see Addis, 2018, 2020, Loftus, 2024, Schacter and Thakral, 2024). Thus, retrieved episodic information may sometimes only partially determine the content of a memory representation (although, as we argued above, the content must have *some* causal connection to a prior event to count as a memory[11] [it is however worth noting that in some circumstances incidentally encoded memories for complex events can remain highly accurate over extended time periods, (e.g., Diamond et al., 2020; see also Brewin et al., 2020)]. Together, these and other factors act to reduce the overlap between encoding- and retrieval-related neural activity to considerably less than the 100 % illustrated in Fig. 2. Intriguingly, findings from studies examining retrieval-related reactivation of engram cells in rodents also suggest that reinstatement is far from complete, with only a minority of the neurons identified as active during an encoding event demonstrating reactivation at retrieval (for review, see Richards and Frankland, 2017).

3.2. Memory consolidation and reinstatement

According to the account given above, a key feature of retrieval-related reinstatement is its dependence on the hippocampus. This is held to be a consequence of the combination of several functional properties of the structure, including its role as a 'hub' where different types of information (e.g. object and spatial) converge, and its efficacy at decorrelating similar input patterns (minimizing overlap between the resulting engrams through pattern separation), and selectively activating engrams in response to a sufficiently specific retrieval cue (pattern completion). A currently open question is whether the reinstatement account should be extended to include 'consolidated' memories that seemingly do not depend on the hippocampus. According to 'standard consolidation theory', for example (Squire et al., 2015), the dependency of an episodic memory on the hippocampus is time-limited and exists only until the memory has become established in the neocortex. If this theory is correct, then the question arises as to what supplants the pattern completion operations initially supported by the hippocampus that allow a retrieval cue to elicit reinstatement of encoded episodic information? One possible answer is that the hippocampus is supplanted by a neocortical region such as the medial PFC, which takes on the 'indexing' role previously played by the hippocampus (Euston et al., 2012). Alternately, it might be that, over time, neocortical-neocortical connections become strengthened sufficiently to permit a retrieval cue to elicit activation of the entirety of a distributed, neocortically housed engram. The question is moot, however, if retrieval of episodic memories always depends on the hippocampus (for review see Moscovitch and Gilboa, 2024; see also Yonelinas et al., 2019).

---

[11] Just how strong that connection needs to be is an interesting and, within the perspective advanced here, unresolved, question.

According to this view, systems consolidation does not preserve an episodic memory but instead 'transforms' it, such that it acquires the character of a semantic memory. In an extension of this idea, Gilboa and Moscovitch (2021) proposed that a semantic memory about a specific event might be formed concurrently with the formation of the corresponding episodic memory, rather than subsequent to it (see Fig. 1 of Gilboa and Moscovitch, 2021). Regardless, the originally encoded episodic memory is held to always depend on the hippocampus, where it can be refreshed or modified if activated by a retrieval cue or, if left dormant, degrade over time and eventually become inaccessible (see Box 1). Thus, episodic memories that 'remain detailed and context-specific' (Moscovitch and Gilboa, 2024, p. 1287) never lose their dependence on the hippocampus. Consequently, if the hippocampal engram encoding an episodic memory becomes inaccessible, memory for specific details of the episode will no longer be available. It should be noted, however, that it has been reported that amnesic individuals suffering extensive hippocampal damage can acquire novel name-object associations in only two study trials when the associations are presented in the context of a 'fast mapping' study procedure (Sharon et al., 2011). This finding implies that there are circumstances in which new memories can be rapidly established in the neocortex independently of the hippocampus. The finding has proven difficult to replicate however, raising concerns about its generality (see Cooper et al., 2019 and associated papers in the same issue). Relatedly, it has been proposed, on the basis of evidence from functional and structural neuroimaging, that repeated study-test cycles can facilitate the rapid formation of what the authors refer to as an 'engram' in medial parietal cortex (Brodt et al., 2018, 2016). It should be noted however that unlike the case of fast mapping described earlier, there is no suggestion that the hippocampus is unnecessary for the establishment of the putative neocortical engram.

What is the role of reinstatement in the retrieval of semantic memories? According to 'sensorimotor' theories, the 'engram' representing a semantic memory is widely distributed across modality-selective neocortical regions and is encoded in the same neural ensembles that respond selectively to the perceptual and motoric features activated when the represented concept is directly experienced (e.g., Chao and Martin, 1999; Simmons et al., 2007). In some models (e.g., Martin, 2016), it is assumed that partial activation of such a neural representation is sufficient for activation to directly spread between different neocortical regions and fully activate the memory representation. In other models, for example, the 'hub and spoke' model of Patterson et al. (2007), retrieval of a memory requires the contribution of one or more neocortical convergence zones (hubs) that store the patterns of co-activation across distributed neural regions that contribute to the neural representation of the retrieved concept. In either case, sensorimotor theories propose that semantic memories are represented by patterns of neural activity that overlap with the patterns elicited when the referents of the memories are directly experienced – that is, as patterns of 'reinstated' neural activity (for further discussion, see Renoult et al., 2019).

3.3. fMRI and reinstatement

Given the dominance of fMRI in studies of memory reinstatement in humans, three additional points are worth making about the approach. First, most studies have examined reinstatement at the 'category' rather than the 'item' level (for a recent review and discussion of this distinction, see Rugg and Srokova, 2024). Category level reinstatement refers to overlap between encoding and retrieval in the activity elicited by exemplars belonging to different classes of study event, for example, events containing face versus scene images (e.g., Hill et al., 2021), or requiring different cognitive operations (e.g., Johnson et al., 2009). Reinstatement at this level is indicative of memory for aspects of a study event that are shared with other members of the same class of event. Hence, the mnemonic content associated with category-level reinstatement is necessarily generic (e.g., 'this retrieval cue was associated with a face rather than a scene', a level of discrimination that is sufficient to support only a relatively undifferentiated memory of the 'gist' of an event). By contrast, item-level reinstatement refers to the reinstatement of the pattern of encoding activity that is idiosyncratic to a specific event, for example, an image of a particular scene. It is this form of reinstatement that likely contributes to memory representations that contain event-specific information. Surprisingly, only a few published studies have examined memory reinstatement at the item level (e.g., Hill et al., 2021; Kuhl and Chun, 2014; Wing et al., 2015).

Second, recent findings indicate that the term 'reinstatement' is something of a misnomer if it is interpreted to mean that encoding- and retrieval-related activity always exactly mirror one another. In at least some circumstances, the anatomical loci of the two classes of activity can differ systematically (for review, see Favila et al., 2020). For example, reinstatement of neural activity elicited by visual scenes has been reported to demonstrate an 'anterior shift' (a term first applied to analogous phenomena in studies of semantic memory, Rugg and Thompson-Schill, 2013), in that the maxima of retrieval-related activity in scene-selective cortical regions such as the parahippocampal cortex is displaced anteriorly by some 5–10 mm with respect to the maxima at encoding (Srokova et al., 2022; Steel et al., 2021). While alternative accounts of the anterior shift have been proposed (Favila et al., 2020; Steel et al., 2021), one of the present authors and his colleagues (Srokova et al., 2022) have suggested that it reflects 'representational re-weighting', such that retrieved information about an event is weighted more heavily towards conceptual information (relative to sensory information) than was the on-line representation of the event (see Deng et al., 2021 for a similar proposal). This account is motivated by the long-standing idea that occipito-temporal cortical regions demonstrate a 'representational gradient', such that more posterior aspects support the representation of lower-level, physical features of a stimulus event, while anterior aspects support more abstract semantic or conceptual representations (e.g., Kravitz et al., 2013; Peelen and Caramazza, 2012; Simmons et al., 2007). The re-weighting account is consistent with the finding that fine-grained perceptual details of an event are retained in memory less well than more general perceptual or conceptual features (e.g., Sacripante et al., 2023; Sekeres et al., 2016; Zeng et al., 2021).

Lastly, although it is tempting to draw a direct parallel between retrieval-related reinstatement as operationalized in human functional imaging studies and studies of reinstatement at the single neuron level (e.g., Josselyn and Tonegawa, 2020; Khuvis et al., 2021), a word of caution is in order. Whereas the neuroimaging findings leave no doubt that encoding-retrieval overlap can be detected at the level of small populations of voxels, the spatial resolution attainable with even the most advanced non-invasive imaging methods is several orders of magnitude too coarse to license the conclusion that the overlap extends to individual neurons.[12] While this might be a reasonable assumption, it is no more than that. If the assumption is valid, however, research in genetically modified mice over the past decade or so might provide useful insights into possible mechanistic bases of hippocampally mediated memory reinstatement as it is reflected in fMRI signals. Findings from this research suggest that memories for single events (almost invariably, a context-shock association) are sparsely encoded in widely distributed neuronal ensembles (engram cells). By virtue of their co-activation during the learning event, these cells become interconnected through Hebbian learning, such that activation of a subset of the cells (especially cells located in the hippocampus, Roy et al., 2022) can lead to activation of additional cells belonging to the ensemble. Extending these observations to human episodic memory, one can speculate that as an event is experienced, a similar Hebbian process

---

[12] The disparity in spatial scales is well captured by the fact that a 2x2x2mm neocortical MRI voxel contains approximately 120,000 neurons (Shapson-Coe et al., 2024).

operates to strengthen connectivity within ensembles of co-active neurons located in different functionally specialized neocortical regions, and between these ensembles and a neuronal ensemble located in the hippocampus. Because of this experience-dependent process of synaptic strengthening, subsequent activation of the hippocampal ensemble will lead to reactivation of neurons belonging to the neocortical ensembles that were active during the original experience, and hence to memory reinstatement.

4. Stability and malleability of memory representations

There are a number of factors that cause memories to change after they have been encoded. Some of these factors apply to both episodic and semantic memory, although much more work has been conducted on the malleability of episodic (e.g., Addis, 2018; Barry and Maguire, 2019; Schacter and Thakral, 2024) than semantic memory (but see Barsalou, 2009; Kumar, 2021; Reilly et al., 2016). As was noted above in Section 2.3 episodic memory has typically been studied, at least in the laboratory, by investigating the correspondence between an input (i.e., a study item) and a later behavioral response (i.e., a memory judgment). This is rarely the case for semantic memory, where, except for a relatively small number of studies investigating learning of new concepts (e. g., Murphy, 2002), memory 'tests' typically involve conceptual knowledge acquired in the remote past.

4.1. Time and experience-dependent variability in memory

4.1.1. Forgetting

Human forgetting can be intentional or unintentional (Fawcett et al., 2024) but here we focus on unintentional forgetting (for recent reviews on intentional forgetting, see Anderson and Hulbert, 2021; Fawcett et al., 2024). Forgetting can be defined as *"the inability to recall something now that could be recalled on an earlier occasion"* (Tulving, 1974). As Tulving implied, this definition does not necessarily mean that the corresponding memory has been permanently lost, but, rather, that it is not currently accessible. Consistent with this possibility, recent findings from non-human animal research suggest that forgetting can indeed reflect a reversible decline in memory accessibility (O'Leary et al., 2024; see also Box 1 and, for relevant human studies, MacLeod and Macrae, 2001 and Storm et al., 2012).

Classical accounts of forgetting, based on forgetting curves (also known as retention functions), seek to explain why forgetting follows a negatively accelerated function over time (Anderson, 2025; Ebbinghaus, 1885; Murre and Dros, 2015; but see Radvansky et al., 2022). Early explanations of this pattern focused on the possible contributions of time-dependent decay and memory interference, the argument in the latter case being that memories become less sensitive to interference with increasing time as a result of a time-dependent consolidation process (Wixted, 2004a, b).

Interference theories of forgetting have found favor with a number of contemporary investigators (e.g., Anderson, 2003; Lechner et al., 1999; Robertson, 2012; Wixted, 2004a, b). For instance, while not rejecting the possibility of passive memory decay, Wixted (2004ab) argued for the crucial role of retroactive interference. Reviewing evidence from psychology, psychopharmacology and neuroscience, he proposed that both the formation of new memories and the consolidation of recently acquired memories rely on the hippocampus, with newly formed memories more vulnerable to interference, and thus to forgetting, than older memories. This is because new memories have yet to be consolidated or are only partially consolidated, rendering them vulnerable to competition with later acquired memories, perhaps because of competition for a 'limited pool of hippocampal resources' (Wixted, 2004b, p. 264). It is assumed that as consolidation proceeds a memory loses its dependence on the hippocampus and becomes neocortically established (the 'standard consolidation theory' described above), and in doing so lessens its vulnerability to interference. Of importance, Wixted argued that the

deleterious effects of interference are generic and not confined to the encoding of new information that resembles or overlaps with previously learned information. Rather, any kind of post-encoding processing has the propensity to interfere with the consolidation of a previously encoded memory.

A radically different interference-based account of forgetting was proposed by Yonelinas et al. (2019), who argued that the concept of consolidation is unnecessary to explain time-dependent forgetting. According to their 'contextual binding' (CB) account, forgetting is largely the result of 'contextual interference'. From this perspective, events that occur shortly before or after a to-be-remembered event will share its context and, consequently, have the potential to interfere with subsequent retrieval of the event by virtue of contextually cued co-activation. An advantage of this account is that, unlike standard consolidation theory, it can explain why episodic memories are negatively impacted not only by the occurrence of post-encoding events, but also by events occurring prior to encoding. A potential weakness of the account, however, is the absence of a clear explanation, beyond an appeal to the effects of repeated retrieval, of how an episodic memory becomes established in the neocortex in a semanticized form (see Section 3.2).

Despite their marked divergences, the two accounts of forgetting outlined above appear to share the assumption that a memory is either wholly retrievable or wholly irretrievable (forgotten). This assumption is highly questionable. As is well documented, different features or aspects of an encoded memory can be forgotten at different rates. For example, and as was noted above (Section 3.3), 'low level' sensory and perceptual features of an event are more likely to be forgotten than central and conceptual features. Thus, a fully developed theory of forgetting will need to include an account of why memories of some kinds of features are more vulnerable to the effects of time than others (e.g., Andermane et al., 2021), and a mechanism by which the features become 'unbound' or 'detached' from the memory trace encoding the event.

4.1.2. Degradation of hippocampal engrams

As was discussed in Section 3.2, proposals such as the Trace Transformation Hypothesis (Moscovitch and Gilboa, 2024; Winocur and Moscovitch, 2011) argue that systems consolidation involves not merely the neocortical consolidation of an episodic memory, but its transformation (semanticization). The hypothesis further proposes that the originally encoded memory trace remains in the hippocampus, where it might be strengthened or modified if activated by a retrieval cue (see below). Thus, the trace transformation hypothesis and its later revisions propose that information about an event is retained in multiple forms. It has been proposed however that, if it remains dormant, the hippocampal trace degrades and eventually become inaccessible (Golbabaei and Frankland, 2025; McClelland et al., 1995; see Box 1). As we noted previously (Section 3.2), this view proposes that the hippocampus is always needed to support a detailed episodic memory and that the trace will likely need to be periodically 'refreshed' or re-encoded (see below for the role of retrieval in this process).

4.1.3. Retrieval-induced memory modification

As alluded to above and emphasized by Tulving (1972), who was articulating an idea that goes back at least to Semon (1904), retrieval from either semantic or episodic memory constitutes an episode in its own right that can potentially be 're-encoded' as a new memory. Subsequently, findings from both non-human animal and human research have been interpreted as evidence that the retrieval of an episodic memory can make the memory susceptible to modification, perhaps by returning it to a labile state and hence to the need for 'reconsolidation' (e.g., Misanin et al., 1968; Nader et al., 2000; see also Antony et al., 2017; Dudai, 2012; Scully et al., 2017). In principle, retrieval could lead to modification of the original memory trace (sometimes referred to as memory updating), the 'strengthening' of the trace without modification, the formation of a new trace (as originally envisaged by Tulving),

or to a scenario in which the original trace co-exists with the re-encoded trace. While much of the evidence seems to be compatible with this latter possibility, whether and when retrieval leads to the formation of a new memory trace, as opposed to the strengthening or modification of the existing trace, is unclear (Golbabaei and Frankland, 2025; Phelps and Hofmann, 2019; Tompary and Murty, 2025). Moreover, the amount of updating of the original trace, and its subsequent accessibility, might vary with the relative relevance of the original and modified memories to current and anticipated circumstances (Wahlheim and Zacks, 2025).

Findings from behavioral studies in humans using 'memory updating' paradigms have been interpreted as evidence that a 'reminder cue' can lead to the addition of new information to an existing memory (Hupbach et al., 2007, 2009, 2008). For example, in the study of Hupbach et al. (2007), participants studied two separate lists of items (A and B). Prior to studying items from list B, some participants were reminded of list A. These participants showed a higher number of list B intrusions than the control group when asked to recall list A items, a finding that was interpreted as indicating that 'reactivation' of a memory allows for its updating (see also Hupbach et al., 2009). Updating effects such as these have been reported to emerge gradually over time and to be specific to the reactivated study list (see also Lee et al., 2017; Scully et al., 2017; St Jacques et al., 2015).

It has been argued however that updating effects like those described above can be explained in terms of the influence of 'temporal context' (Sederberg et al., 2011) rather than as a consequence of reconsolidation. According to this explanation, in the reminder condition, the reactivated list A context serves as the context for list B, thus linking both lists of items to a common context. Subsequent attempts to recall list A items will 'reactivate' this shared context, facilitating access to items belonging to both lists and leading to list B intrusions. From this perspective, therefore, the presence of list B intrusions does not signify modification of the original List A memories.

In a similar vein, studies employing the 'misinformation paradigm' (Loftus and Hoffman, 1989) demonstrate that retrieval of the memory of an event can make participants susceptible to inaccurate information about it (e.g., Butler and Loftus, 2018; Chan et al., 2009; Rindal et al., 2016). In this paradigm, participants are first exposed to a complex event, such as a brief film (Loftus, 1975; Loftus and Palmer, 1974). Subsequently, they receive erroneous information about the event, after which their memory is tested. A consistent finding is that participants tend to incorporate the erroneous information into their reported memory of the original event. Such findings have frequently been interpreted as reflecting a modification of the originally encoded memory (e.g., Barry and Maguire, 2019; Loftus and Loftus, 1980). However, like the 'reconsolidation' findings described earlier, the misinformation effect can also be explained by temporal context models. According to this account, erroneous memory judgments are a consequence of the attribution of the post-event information to the wrong source (Sederberg et al., 2011; see Johnson et al.,1993 for a similar interpretation). This account leaves the original memory intact and, given a sufficiently specific cue, accessible (e.g., Mccloskey and Zaragoza, 1985).

Somewhat in keeping with the foregoing account, recent neuroimaging findings from the misinformation paradigm have been interpreted according to a 'multiple trace' model. According to this model, the hippocampal trace representing the original event remains intact and competes for expression with the memory trace of the misinformation (Shao et al., 2023). The basis for this claim was the finding that across-voxel patterns of activity in the hippocampus elicited by the original experience and the misinformation appear to co-exist. Shao et al. (2023) proposed that source monitoring processes supported by lateral prefrontal regions play a key role in resolving the conflict between the two classes of mnemonic information (see also Brunswick et al., 2025).

Findings from other neuroimaging studies are also consistent with the possibility of a coexistence of an original memory trace and a newer

trace associated with the same items. In an fMRI study by Kuhl et al. (2012), participants repeatedly studied words, either associated with the same picture (e.g., a face) or with a different picture (e.g., an object) before being asked to retrieve the most recent associated picture with the word as cue. The authors trained a pattern classifier to discriminate between the picture categories (faces, objects, but also scenes) and reported parallel activation in ventral temporal cortex for new memories (i.e., the object recently associated with the word) and older memories (i.e., the face originally associated with the word), even when newer memories were successfully retrieved by the participants. Stronger activation of older memories was associated with slower reaction time and with activation of the anterior cingulate cortex, while activity in the left inferior frontal gyrus predicted successful retrieval of newer memories.

Finally, the findings from another recent neuroimaging study (Molitor et al., 2021) employing multi-voxel pattern analyses suggest that reactivation of a memory during new learning can lead to subsequent expression of a memory representation that combines the original and the new information and, concurrently, to the expression of separate representations. An AB-AC paradigm was employed in which two images (B and C) became indirectly related to one another through their common association with image A. When memory was later cued with image A, memory differentiation (distinct A-B and A-C activity patterns) was identified in the dentate gyrus/CA2,3 subfields and the subiculum, whereas memory integration (overlap between the two patterns) was evident in the CA1 sub-field. For the reasons outlined in Box 2, however, findings such as these do not necessarily imply modification at the level of the underlying engrams.

As was noted by Roediger and Abel (2022), memory retrieval can be viewed as a 'double-edged sword': while retrieval can enhance memory for items that are correctly recalled, it can also strengthen 'memories' for incorrectly recalled material (whether erroneously recalled by participants spontaneously or induced experimentally via exposure to misinformation). In addition to these retrieval-induced increases in correct and incorrect recall, memory retrieval can also lead to the weakening of competing memories (e.g., items associated with the same retrieval target), a phenomenon referred to as retrieval-induced forgetting and that is thought to reduce interference via an inhibitory control mechanism (Anderson et al., 1994; Wimber et al., 2015). Moreover, as was noted by Nadel and Moscovitch (1997, p. 223), 'each reactivation of a memory trace occurs in an altered neuronal and experiential context'. Retrieval thus allows a memory to become associated with a new context: in such cases, later attempts to retrieve the episode will benefit from reinstatement not just of the original encoding context (e.g., Xue et al., 2010), but the context associated with its 're-retrieval' (for reviews, see Roediger and Abel, 2022; Gilboa and Moscovitch, 2021).

Together, the evidence from human behavioral and neuroimaging studies is compatible with the proposal that memory retrieval can promote the co-existence of an original memory trace and a new trace. Thus, the seeming incorporation of new information into an existing memory does not necessarily mean that the original memory is no longer available (Phelps and Hofmann, 2019). An important implication of these findings is that it is unclear whether there is such a thing as 'memory updating', that is, retrieval-induced modification of the content of a reactivated memory trace. Rather, the original memory may simply be 'out competed' by a newly formed one. This conception of memory updating is consistent with the results from a recent study in rodents in which it was reported that reminder cues initiated the formation of a *de novo* hippocampal ensemble of engram cells which demonstrated little or no overlap with the ensemble associated with the originally encoded memory (Lei et al., 2025).

4.2. Memory construction

The content of a retrieved episodic memory is commonly proposed to reflect the outcome of the same constructive processes that are engaged

during other types of mental construction, such as future thinking or imagination. A major line of evidence held to support this contention is that episodic retrieval, future thinking and imagination all engage the same or a very similar large scale brain network, which has variously been referred to as the 'default mode', autobiographical', 'core recollection' or just the 'core' network (Addis, 2018; Rugg and Vilberg, 2013; Schacter et al., 2007). As has been discussed previously, however (Binder and Desai, 2011; Renoult et al., 2019), essentially the same network (now in the guise of the 'semantic network'; Binder and Desai, 2011) is also engaged during *semantic* memory retrieval. Therefore, it would seem either that semantic retrieval engages the same constructive processes argued to be common to episodic retrieval, imagination and future thought, or that the explanation for the overlap in the activation patterns observed across these different cognitive activities lies elsewhere. Following Renoult et al. (2019), we propose that the common factor is their reliance on the extensive engagement of semantic (conceptual) processing and not a shared mental construction process (see also Robins, 2022).

Regardless of this issue, an important aspect of the constructive perspective is that the content of a memory representation is determined, at least in part, by the retrieval process itself, rather than merely reflecting the invariant content of a 'reactivated' memory trace (e.g., Klein, 2013; Moscovitch, 2007). For example, to retrieve a specific memory (e.g., the New Year party you attended in 2023), you might rely on general knowledge about the people likely to have been present, the likely location, the type of event, and so on (e.g., Reagh and Ranganath, 2018; Sekeres et al., 2018). Such constructivist views of memory imply that the content of a memory for the same event might differ each time it is retrieved in ways that distance the memory from the initial experience. The addition of such generic information likely depends on the form of the memory and the integrity of the relevant hippocampal engram. For example, compared with traditional 'list learning' laboratory experiments, the complexity of autobiographical memory provides many more opportunities for the influence of extra-mnemonic information (Wardell and Palombo, 2024), especially for remote memories when the hippocampal engram may have become severely degraded or inaccessible. In such cases, verification of a causal link between a reported memory and the event it purports to correspond with can be challenging and may rely largely on the confidence with which the rememberer asserts that the event occurred (Johnson et al., 1988; Rubin et al., 2003).

As we discussed above, sensorimotor theories propose that semantic memories are represented by distributed patterns of neural activity that overlap with the patterns elicited when the referents of the memories were directly experienced. However, as for episodic memory, functional neuroimaging research suggests that, depending on the retrieval context, different patterns of brain activity can be associated with the retrieval of ostensibly the same concept (Reilly et al., 2016, 2025). For example, motor regions are activated when an object is processed with respect to potential actions, but not when it is processed in respect of its color (van Dam et al., 2012; see also Raposo et al., 2009; Willems and Casasanto, 2011). Thus, the features that are incorporated into an active representation of the concept are not invariant but differ according to task and contextual factors (Kumar, 2021). As has been proposed in the case of episodic memory (Rugg, 2024), selection of the goal-relevant features of a concept has been argued to depend on top-down control processes (e.g., Badre and Wagner, 2007; Lambon Ralph et al., 2017; Novick et al., 2009; Vatansever et al., 2021). A constructive perspective of semantic memory might be compatible with some contemporary 'embodied' perspectives, by which semantic processing takes the form of 'simulation' processes that are driven by pattern completion mechanisms (Barsalou, 2009). Such approaches imply that knowledge of concepts is grounded in the brain's sensory and motor systems, as described previously (see Section 3.2), and that access to this knowledge involves a simulation of the sensorimotor processing engaged by the referent of the concept. According to this view, simulation is a basic computational mechanism that applies to all kinds of memory (Barsalou, 2003; see also Addis, 2018, 2020). Pattern completion in this context refers to unconscious 'inferential' processes based on previous knowledge. For example, upon seeing a new coffee house, one would 'infer' 'that it contains tables, chairs, and people talking, reading and working' (Barsalou, 2016). According to this framework, the level of detail in a simulation can vary widely, from highly detailed to skeletal and vague, depending on current goals and task demands (Barsalou, 2009).

Thus, while the evidence is more limited than in the case of episodic memory, the literature on semantic memory is compatible with the idea that, like episodic memory, the manner in which a semantic memory 'trace' influences the content of an active semantic memory representation is context dependent. That is, the content of the representation can be influenced by external and internal factors just as in the case of episodic memory. And, also as proposed for episodic memory, these factors might be more relevant for complex (e.g., knowledge about events) than more basic semantic knowledge (e.g., knowledge about category exemplars). Finally, one can speculate that the extent to which a semantic memory trace is updated will vary with the relevance of existing semantic knowledge to current and anticipated future situations, and whether the knowledge is obsolete (see Wahlheim and Zacks, 2025).

4.3. Concluding comments

We hope that, like ourselves, readers will find value in thinking about the cognitive neuroscience of memory from a representational perspective. We particularly hope such a perspective will be helpful in clarifying unresolved empirical and theorical issues and considering new research directions. In developing a cognitive neuroscientific account of human episodic and semantic memory that is informed by ideas about neural and cognitive representations, we have benefited greatly from the rich philosophical literature on this topic, as well as the recent literature describing research focusing on 'engram cells'. At the least, we hope that our appeal for terminological and conceptual clarity in the use of the term 'representation' in the cognitive neuroscience of memory – especially regarding the benefits of distinguishing between neural versus cognitive, and active versus latent representations - will strike a chord with other researchers in the field.

5. Box 1

5.1. Mechanisms of change in memory representations: cellular and molecular findings

One factor often discussed in relation to degradation of memory traces is molecular turnaround. In mammals, most excitatory axodendritic synapses are found on dendritic spines (Grutzendler et al., 2002). In the hippocampus, significant amounts of spine turnover may occur within 1–2 weeks, with Attardo et al. (2015) estimating 100 % turnover in 3–6 weeks in mice. These observations have been interpreted by some investigators as a significant challenge for the long-term maintenance of hippocampal engrams and an important reason they would be expected to weaken over time (e.g. Barry and Maguire, 2019). However, as noted by Moscovitch and Nadel (2019), cellular instabilities such as those observed in the hippocampus would not necessarily prevent a neural ensemble as a whole from stably encoding long-lasting engrams. This view is consistent with the non-human animal literature on 'representational drift' that suggests that within a neuronal ensemble, different neurons are responsible for encoding a memory at different times (Rule et al., 2019). Representational or neural drift (Lopez et al., 2024) refers to the observation that, in various neocortical areas and in the hippocampus, neurons tend to change their tuning properties over time. Thus, persistence of representations is maintained at the population rather than the single cell level (Keinath et al., 2022; Lopez et al., 2024). Moreover, in addition to synaptic
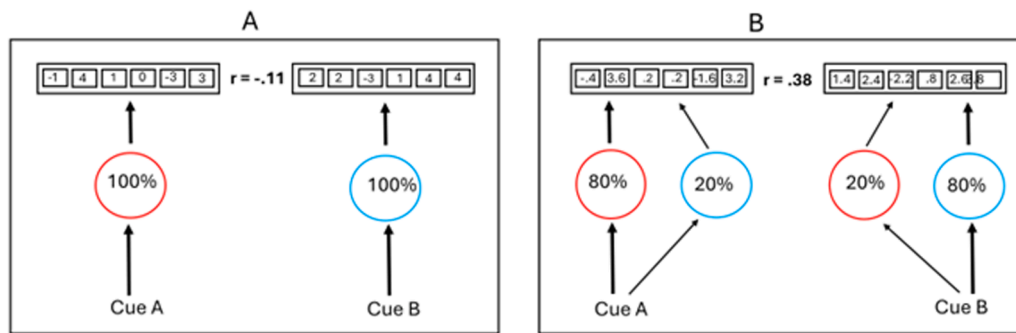
Fig. 3. Cartoon depiction of how the similarity of across-voxel patterns of memory-related fMRI BOLD activity can change without modification of the underlying engrams. Top rows depict BOLD signal amplitudes across six voxels belonging to a region of interest; middle rows depict two different engrams (red and blue circles), and the percentages indicate their relative activation in response to two retrieval cues (A and B). A: The engrams are activated exclusively by the respective cues, leading to the expression of dissimilar activity patterns (Pearson's r = -.11). B: As a result of some experimental manipulation, each engram remains strongly activated by one of the cues but is also weakly activated by the other cue. The BOLD signal amplitudes in each voxel are now the weighted averages of the amplitudes depicted in A and the similarity between the resulting activity patterns increases (Pearson's r = .38). Thus, active memory representations can become 'integrated' while their respective engrams remain independent.

plasticity such as that arising from modification of dendritic spines, intra-cellular modifications (such as RNA, DNA methylation and histone modification) could also be crucial to the long-term maintenance of engrams (Gallistel, 2017, 2021; Gershman, 2023), serving as the "engram code" (Gallistel, 2017). For instance, DNA methylation has been reported to be necessary for the formation of neocortically dependent remote memories (Tonegawa et al., 2018).

Finally, the non-human animal literature has described so-called 'silent' engrams which cannot be activated by natural cues, but which can be activated artificially (e.g., through optogenetic stimulation; Tonegawa et al., 2018). Such phenomena are reminiscent of Tulving's distinction between memory availability and accessibility (silent engrams representing memories that are available but not accessible; Tulving and Pearlstone, 1966).

6. Box 2

6.1. fMRI multi-voxel pattern analyses and memory representations

The development of multi-voxel pattern analysis (MVPA) – a collection of fMRI analysis approaches focused on across-voxel profiles of blood oxygen level dependent (BOLD) signal amplitude – has led to the use of fMRI to examine 'representational' questions in cognitive neuroscience (indeed, the name of one such analysis approach is 'Representational Similarity Analysis', Kriegeskorte et al., 2008). The idea that MVPA can provide insights into neural representations arises from the assumption that the representations are manifest in patterns of neural activity that are distributed across spatial scales compatible with the spatial resolution of fMRI. The underlying premise is that the more similar two neural patterns are to one another, the more similar are their neural representations. This premise is at the heart of the employment of MVPA to examine phenomena like retrieval-related reinstatement (e.g. Johnson et al., 2009), and the factors responsible for memory 'integration' versus memory 'separation' (e.g. Schlichting and Preston, 2015), to give just two examples.

There is no doubt that MVPA has a valuable role in addressing questions about the variables responsible for similarities (and dissimilarities) in patterns of neural activity elicited by different experimental items or in different experimental conditions. Nor is there any doubt that MVPA can yield theoretical insights that go beyond what can be achieved with 'univariate' approaches to fMRI data analysis. Nonetheless, two caveats are worth noting. First, while it is tempting to interpret variability in across-voxel profiles of BOLD activity in terms of variability in neural representations, such interpretations are not necessarily valid. Davis et al. (2014) compellingly demonstrated that differences in

the similarity of multi-voxel activity patterns can be driven entirely by univariate effects, that is, by differences in the across-voxel mean BOLD signal, and hence differences between patterns may have no representational implications (see also Gessell et al., 2021). Second, even when MVPA findings can be given a representational interpretation, the findings are relevant only at the level of *active* representations (note that such representations could be active at the neural level only, with no corresponding cognitive representation, as might be the case when memories are cued during sleep; see Section 2.2), and only as these representations are expressed in the brain at a relatively large spatial scale (see footnote 12). Thus, the finding that a manipulation results in the multi-voxel activity profiles associated with the retrieval of two memories becoming more similar to one another (becoming 'integrated') does not necessarily mean that the memory traces representing the memories have been modified or integrated. Instead, the findings might reflect changes in the relative strengths of simultaneous activation of the different traces, leading to a change in the 'blending' of the neural patterns associated with the separate activation of each trace (see Fig. 3).

Data availability

No data was used for the research described in the article.

References

Addis, D.R., 2018. Are episodic memories special? On the sameness of remembered and imagined event simulation. J. R. Soc. N. Zeal. 48, 64–88.

Addis, D.R., 2020. Mental time travel? A neurocognitive model of event simulation. Rev. Philos. Psychol. 11, 233–259.

Alvarez, P., Squire, L.R., 1994. Memory consolidation and the medial temporal lobe: a simple network model. Proc. Natl. Acad. Sci. USA 91, 7041–7045.

Andermane, N., Joensen, B.H., Horner, A.J., 2021. Forgetting across a hierarchy of episodic representations. Curr. Opin. Neurobiol. 67, 50–57.

Anderson, M.C., 2003. Rethinking interference theory: executive control and the mechanisms of forgetting. J. Mem. Lang. 49, 415–445.

Anderson, M.C., 2025. Incidental Forgetting. In: Baddeley, A., Eysenck, M.W., Anderson, M.C. (Eds.), Memory, 4th edition ed. Routledge.

Anderson, M.C., Bjork, R.A., Bjork, E.L., 1994. Remembering can cause forgetting - retrieval dynamics in Long-Term-Memory. J. Exp. Psychol. Learn. Mem. Cogn. 20, 1063–1087.

Anderson, M.C., Hulbert, J.C., 2021. Active forgetting: adaptation of memory by prefrontal control. Annu. Rev. Psychol. 72 72, 1–36.

Andonovski, N., in press. Memory and Causation, in: Sant'Anna, A., Craver, C.F. (Eds.), The Oxford Handbook of Philosophy of Memory. Oxford University Press.

Antony, J.W., Ferreira, C.S., Norman, K.A., Wimber, M., 2017. Retrieval as a fast route to memory consolidation. Trends Cogn. Sci. 21, 573–576.

Attardo, A., Fitzgerald, J.E., Schnitzer, M.J., 2015. Impermanence of dendritic spines in live adult CA1 hippocampus. Nature 523, 592.

Badre, D., Wagner, A.D., 2007. Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia 45, 2883–2901.

Baker, B., Lansdell, B., Kording, K.P., 2022. Three aspects of representation in neuroscience. Trends Cogn. Sci. 26, 942–958.

Barry, D.N., Maguire, E.A., 2019. Remote memory and the hippocampus: a constructive critique. Trends Cogn. Sci. 23, 128–142.

Barsalou, L.W., 2003. Situated simulation in the human conceptual system. Lang. Cogn. Process. 18, 513–562.

Barsalou, L.W., 2009. Simulation, situated conceptualization, and prediction. Philos. Trans. R. Soc. B Biol. Sci. 364, 1281–1289.

Barsalou, L.W., 2016. Situated conceptualization offers a theoretical account of social priming. Curr. Opin. Psychol. 12, 6–11.

Bernstein, D.M., Loftus, E.F., 2009. How to tell if a particular memory is true or false. Perspect. Psychol. Sci. 4, 370–374.

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cogn. Sci. 15, 527–536.

Bransford, J.D., Franks, J.J., 1971. Abstraction of linguistic ideas. Cogn. Psychol. 2, 331–350.

Brentano, F., 1874. Psychology from an Empirical Standpoint. Routledge, Abingdon, Oxon.

Brewer, W.F., 1986. Schemas versus mental models in human-memory. Bull. Psychon. Soc. 24, 343. –343.

Brewer, W.F., Treyens, J.C., 1981. Role of schemata in memory for places. Cogn. Psychol. 13, 207–230.

Brewin, C.R., Andrews, B., Mickes, L., 2020. Regaining consensus on the reliability of memory. Curr. Dir. Psychol. Sci. 29, 121–125.

Brodt, S., Gais, S., Beck, J., Erb, M., Scheffler, K., Schönauer, M., 2018. Fast track to the neocortex: a memory engram in the posterior parietal cortex. Science 362, 1045.

Brodt, S., Pöhlchen, D., Flanagin, V.L., Glasauer, S., Gais, S., Schönauer, M., 2016. Rapid and independent memory formation in the parietal cortex. Proc. Natl. Acad. Sci. USA 113, 13251–13256.

Brunswick, C.A., Carpenter, C.M., Dennis, N.A., Kwapis, J.L., 2025. Not the same as it ever was: a review of memory modification, updating, and distortion in humans and rodents. Neurosci. Biobehav. Rev. 174, 106195.

Butler, B.J., Loftus, E.F., 2018. Discrepancy detection in the retrieval-enhanced suggestibility paradigm. Memory 26, 483–492.

Chan, J.C., Thomas, A.K., Bulevich, J.B., 2009. Recalling a witnessed event increases eyewitness suggestibility: the reversed testing effect. Psychol. Sci. 20, 66–73.

Chao, L.L., Martin, A., 1999. Cortical regions associated with perceiving, naming, and knowing about colors. J. Cogn. Neurosci. 11, 25–35.

Chen, J., Leong, Y.C., Honey, C.J., Yong, C.H., Norman, K.A., Hasson, U., 2017. Shared memories reveal shared structure in neural activity across individuals. Nat. Neurosci. 20, 115–125.

Chun, M.M., Turk-Browne, N.B., 2007. Interactions between attention and memory. Curr. Opin. Neurobiol. 17, 177–184.

Conway, M.A., 2009. Episodic memories. Neuropsychologia 47, 2305–2313.

Cooper, E., Greve, A., Henson, R.N., 2019. Little evidence for fast mapping (FM) in adults: a review and discussion. Cogn. Neurosci. 10, 196–209.

Craik, F.I.M., Lockhart, R.S., 1972. Levels of processing - framework for memory research. J. Verbal Learn. Verbal Behav. 11, 671–684.

Danker, J.F., Anderson, J.R., 2010. The ghosts of brain states past: remembering reactivates the brain regions engaged during encoding. Psychol. Bull. 136, 87–102.

Davis, T., LaRocque, K.F., Mumford, J.A., Norman, K.A., Wagner, A.D., Poldrack, R.A., 2014. What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. Neuroimage 97, 271–283.

De Brigard, F., 2023. Memory and Remembering. Cambridge University Press.

de Chastelaine, M., Srokova, S., Monier, S., Olivier, J.M., Rugg, M.D., 2025. Effects of age on the strategic control of recollected content as reflected by modulation of neural correlates of scene retrieval. Neurobiol. Aging 154, 1–15.

Debus, D., 2017. Memory Causation. The Routledge Handbook of Philosophy of Memory. Routledge.

Deng, L., Davis, S.W., Monge, Z.A., Wing, E.A., Geib, B.R., Raghunandan, A., Cabeza, R., 2021. Age-related dedifferentiation and hyperdifferentiation of perceptual and mnemonic representations. Neurobiol. Aging 106, 55–67.

Diamond, N.B., Armson, M.J., Levine, B., 2020. The truth is out there: accuracy in recall of verifiable Real-World events. Psychol. Sci. 31, 1544–1556.

Dretske, F., 2003. Experience as representation (Qualia). Nous 67–82.

Dudai, Y., 2007. Memory, It's all about representations. In: Roediger, H.L.I., Dudai, Y., Fitzpatrick, S.M. (Eds.), Science of Memory: Concepts. Oxford University Press.

Dudai, Y., 2012. The restless engram: consolidations never end. Annu Rev. Neurosci. 35, 227–247.

Ebbinghaus, H., 1885. Memory. A Contribution to Experimental Psychology, 1913 edition. Teachers College, Columbia University, New York.

Elward, R.L., Rugg, M.D., 2015. Retrieval goal modulates memory for context. J. Cogn. Neurosci. 27, 2529–2540.

Euston, D.R., Gruber, A.J., McNaughton, B.L., 2012. The role of medial prefrontal cortex in memory and decision making. Neuron 76, 1057–1070.

Favila, S.E., Lee, H., Kuhl, B.A., 2020. Transforming the concept of memory reactivation. Trends Neurosci. 43, 939–950.

Fawcett, J.M., Taylor, T.L., Megla, E., Maxcey, A.M., 2024. Active intentional and unintentional forgetting in the laboratory and everyday life. Nat. Rev. Psychol. 3, 652–664.

Frankland, P.W., Josselyn, S.A., Köhler, S., 2024. Engrams. Curr. Biol. 34, R559–R561.

Gallistel, C.R., 2017. The coding question. Trends Cogn. Sci. 21, 498–508.

Gallistel, C.R., 2021. The physical basis of memory. Cognition 213.

Gallow, J.D., 2022. The metaphysics of causation, in: Nodelman, E.N.Z.U. (Ed.), The Stanford Encyclopedia of Philosophy.

Gershman, S.J., 2023. The molecular memory code and synaptic plasticity: a synthesis. Biosystems 224.

Gessell, B., Geib, B., De Brigard, F., 2021. Multivariate pattern analysis and the search for neural representations. Synthese 199, 12869–12889.

Gilboa, A., Moscovitch, M., 2021. No consolidation without representation: correspondence between neural and psychological representations in recent and remote memory. Neuron 109, 2239–2255.

Golbabaei, A., Frankland, P.W., 2025. The post-"standard model" age: updating theories of systems consolidation. Neuron 113, 339–341.

Grutzendler, J., Kasthuri, N., Gan, W.B., 2002. Long-term dendritic spine stability in the adult cortex. Nature 420, 812–816.

Guest, O., Martin, A., 2025. Are neurocognitive representations 'small cakes'? PhilSci Archive.

Hill, P.F., King, D.R., Rugg, M.D., 2021. Age differences in Retrieval-Related reinstatement reflect Age-Related dedifferentiation at encoding. Cereb. Cortex 31, 106–122.

Hu, X.Q., Cheng, L.Y., Chiu, M.H., Paller, K., 2020. Promoting memory consolidation during sleep: a Meta-Analysis of targeted memory reactivation. Psychol. Bull. 146, 218–244.

Hupbach, A., Gomez, R., Nadel, L., 2009. Episodic memory reconsolidation: updating or source confusion? Memory 17, 502–510.

Hupbach, A., Gomez, R., Hardt, O., Nadel, L., 2007. Reconsolidation of episodic memories: a subtle reminder triggers integration of new information. Learn. Mem. 14, 47–53.

Hupbach, A., Hardt, O., Gomez, R., Nadel, L., 2008. The dynamics of memory: context-dependent updating. Learn. Mem. 15, 574–579.

Ingram, K.M., Mickes, L., Wixted, J.T., 2012. Recollection can be weak and familiarity can be strong. J. Exp. Psychol. Learn. Mem. Cogn. 38, 325–339.

Irish, M., 2019. On the interaction between episodic and semantic representations – constructing a unified account of imagination. In: Abraham, A. (Ed.), The Cambridge Handbook of Imagination. Cambridge University Press.

James, W., 1914. Talks to Teachers on Psychology: And to Students on Some of Life's Ideals. H. Holt and company, New York.

Jeunehomme, O., D'Argembeau, A., 2019. The time to remember: temporal compression and duration judgements in memory for real-life events. Q J. Exp. Psychol. 72, 930–942.

Jeunehomme, O., D'Argembeau, A., 2020. Event segmentation and the temporal compression of experience in episodic memory. Psychol. Res. Psychol. Forsch. 84, 481–490.

Johnson, M.K., Foley, M.A., Suengas, A.G., Raye, C.L., 1988. Phenomenal characteristics of memories for perceived and imagined autobiographical events. J. Exp. Psychol. Gen. 117, 371–376.

Johnson, M.K., Hashtroudi, S., Lindsay, D.S., 1993. Source monitoring. Psychol. Bull. 114, 3–28.

Johnson, J.D., McDuff, S.G., Rugg, M.D., Norman, K.A., 2009. Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. Neuron 63, 697–708.

Josselyn, S.A., Tonegawa, S., 2020. Memory engrams: recalling the past and imagining the future. Science 367, 39–+.

Keinath, A.T., Mosser, C.A., Brandon, M.P., 2022. The representation of context in mouse hippocampus is preserved despite neural drift. Nat. Commun. 13.

Khuvis, S., Yeagle, E.M., Norman, Y., Grossman, S., Malach, R., Mehta, A.D., 2021. Face-Selective units in human ventral temporal cortex reactivate during free recall. J. Neurosci. 41, 3386–3399.

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. J. Neurosci. 35, 1763–1772.

Klein, S.B., 2013. Making the case that episodic recollection is attributable to operations occurring at retrieval rather than to content stored in a dedicated subsystem of long-term memory. Front. Behav. Neurosci. 7.

Kolers, P.A., 1973. Remembering operations. Mem. Cogn. 1, 347–355.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., Mishkin, M., 2013. The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends Cogn. Sci. 17, 26–49.

Kriegeskorte, N., Diedrichsen, J., 2019. Peeling the onion of brain representations. Annu. Rev. Neurosci. 42 (42), 407–432.

Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2, 4.

Kuhl, B.A., Bainbridge, W.A., Chun, M.M., 2012. Neural Reactivation Reveals Mechanisms for Updating Memory. J. Neurosci. 32, 3453–3461.

Kuhl, B.A., Chun, M.M., 2014. Successful remembering elicits event-specific activity patterns in lateral parietal cortex. J. Neurosci. 34, 8051–8060.

Kumar, A.A., 2021. Semantic memory: a review of methods, models, and current challenges. Psychon. Bull. Rev. 28, 40–80.

Lambon Ralph, M.A., Jefferies, E., Patterson, K., Rogers, T.T., 2017. The neural and computational bases of semantic cognition. Nat. Rev. Neurosci. 18, 42–55.

Lechner, H.A., Squire, L.R., Byrne, J.H., 1999. 100 years of consolidation–remembering muller and pilzecker. Learn Mem. 6, 77–87.

Lee, J.L.C., Nader, K., Schiller, D., 2017. An update on memory reconsolidation updating. Trends Cogn. Sci. 21, 531–545.

Lei, B., Kang, B.L., Hao, Y.J., Yang, H.Y., Zhong, Z.H., Zhai, Z.H., Zhong, Y., 2025. Reconstructing a new hippocampal engram for systems reconsolidation and remote memory updating. Neuron 113.

Lewis, D., 1973. Causation. J. Philos. 70, 556–567.

Loftus, E.F., 1975. Leading questions and eyewitness report. Cogn. Psychol. 7, 560–572.

Loftus, E.F., 2024. Memory, but not memory. P Am. Philos. Soc. 165.

Loftus, E.F., Hoffman, H.G., 1989. Misinformation and memory: the creation of new memories. J. Exp. Psychol. Gen. 118, 100–104.

Loftus, E.F., Loftus, G.R., 1980. On the permanence of stored information in the human-brain. Am. Psychol. 35, 409–420.

Loftus, E.F., Palmer, J.C., 1974. Reconstruction of automobile destruction - example of interaction between language and memory. J. Verbal Learn. Verbal Behav. 13, 585–589.

Lopez, M.R., Wasberg, S.M.H., Gagliardi, C.M., Normandin, M.E., Muzzio, I.A., 2024. Mystery of the memory engram: history, current knowledge, and unanswered questions. Neurosci. Biobehav. Rev. 159.

MacLeod, M.D., Macrae, C.N., 2001. Gone but not forgotten: the transient nature of retrieval-induced forgetting. Psychol. Sci. 12, 148–152.

Marr, D., 1971. Simple memory: a theory for archicortex. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 262, 23–81.

Martin, A., 2016. GRAPES-Grounding representations in action, perception, and emotion systems: how object properties and categories are represented in the human brain. Psychon. Bull. Rev. 23, 979–990.

Martin, C.B., Deutscher, M., 1966. Remembering. Philos. Rev. 75, 161–196.

McClelland, J.L., McNaughton, B.L., Oreilly, 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. Psychol. Rev. 102, 419–457.

Mccloskey, M., Zaragoza, M., 1985. Misleading postevent information and memory for events - arguments and evidence against memory impairment hypotheses. J. Exp. Psychol. Gen. 114, 1–16.

Menzies, P., Beebee, H., 2025. In: Zalta, E.N.N., Uri (Eds.), Counterfactual Theories of Causation. The Stanford Encyclopedia of Philosophy.

Michaelian, K., 2016. Episodic Memory and Our Knowledge of the Personal Past. The MIT Press.

Michaelian, K., Robins, S.K., 2018. Beyond the causal theory? Fifty years after martin and deutscher. In: Michaelian, K., Debus, D., Perrin, D. (Eds.), New Directions in the Philosophy of Memory. Routledge, New York, pp. 13–32.

Michaelian, K., Sutton, J., 2017. Memory, in: Zalta, E.N. (Ed.), The Stanford Encyclopedia of Philosophy.

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, 554–555.

Molitor, R.J., Sherrill, K.R., Morton, N.W., Miller, A.A., Preston, A.R., 2021. Memory reactivation during learning simultaneously promotes dentate Gyrus/CA pattern differentiation and CA memory integration. J. Neurosci. 41, 726–738.

Morris, C.D., Bransford, J.D., Franks, J.J., 1977. Levels of processing versus transfer appropriate processing. J. Verbal Learn. Verbal Behav. 16, 519–533.

Moscovitch, M., 2007. Why the engram is elusive. In: Roediger, H.L., Dudai, Y., Fitzpatrick, S.M. (Eds.), Science of Memory: Concepts. Oxford university press, pp. 17–21.

Moscovitch, M., Gilboa, A., 2024. In: Kahana, M.J., Wagner, A.D. (Eds.), Systems Consolidation, Transformation, and Reorganization: Multiple Trace Theory, Trace Transformation Theory, and Their Competitors. Handbook on Human Memory. Oxford University Press.

Moscovitch, M., Nadel, L., 2019. Sculpting remote memory: enduring hippocampal traces and vmPFC reconstructive processes. Trends Cogn. Sci. 23, 634–635.

Murphy, G.L., 2002. The Big Book of Concepts. MIT Press, Cambridge, Mass.

Murre, J.M.J., Dros, J., 2015. Replication and analysis of ebbinghaus' forgetting curve. PLOS One 10.

Nadel, L., Moscovitch, M., 1997. Memory consolidation, retrograde amnesia and the hippocampal complex. Curr. Opin. Neurobiol. 7, 217–227.

Nader, K., Schafe, G.E., Le Doux, J.E., 2000. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature 406, 722–726.

Norman, K.A., 2010. How hippocampus and cortex contribute to recognition memory: revisiting the complementary learning systems model. Hippocampus 20, 1217–1227.

Novick, J.M., Kan, I.P., Trueswell, J.C., Thompson-Schill, S.L., 2009. A case for conflict across multiple domains: memory and language impairments following damage to ventrolateral prefrontal cortex. Cogn. Neuropsychol. 26, 527–567.

Oedekoven, C.S.H., Keidel, J.L., Berens, S.C., Bird, C.M., 2017. Reinstatement of memory representations for lifelike events over the course of a week. Sci. Rep. 7, 14305.

O'Leary, J.D., Bruckner, R., Autore, L., Ryan, T.J., 2024. Natural forgetting reversibly modulates engram expression. eLife 12.

Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat. Rev. Neurosci. 8, 976–987.

Peelen, M.V., Caramazza, A., 2012. Conceptual object representations in human anterior temporal cortex. J. Neurosci. 32, 15728–15736.

Phelps, E.A., Hofmann, S.G., 2019. Memory editing from science fiction to clinical practice. Nature 572, 43–50.

Pitt, D., 2022. In: Zalta, E.N., Nodelman, U. (Eds.), Mental Representation. The Stanford Encyclopedia of Philosophy.

Poldrack, R.A., 2021. The physics of representation. Synthese 199, 1307–1325.

Radvansky, G.A., Doolen, A.C., Pettijohn, K.A., Ritchey, M., 2022. A new look at memory retention and forgetting. J. Exp. Psychol. Learn. Mem. Cogn.

Raposo, A., Moss, H.E., Stamatakis, E.A., Tyler, L.K., 2009. Modulation of motor and premotor cortices by actions, action words and action sentences. Neuropsychologia 47, 388–396.

Reagh, Z.M., Ranganath, C., 2018. What does the functional organization of cortico-hippocampal networks tell us about the functional organization of memory? Neurosci. Lett. 680, 69–76.

Reagh, Z.M., Ranganath, C., 2023. Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. Nat. Commun. 14, 1279.

Reilly, J., Peelle, J.E., Garcia, A., Crutch, S.J., 2016. Linking somatic and symbolic representation in semantic memory: the dynamic multilevel reactivation framework. Psychon. Bull. Rev. 23, 1002–1014.

Reilly, J., Shain, C., Borghesani, V., Kuhnke, P., Vigliocco, G., Peelle, J.E., Mahon, B.Z., Buxbaum, L.J., Majid, A., Brysbaert, M., Borghi, A.M., De Deyne, S., Dove, G., Papeo, L., Pexman, P.M., Poeppel, D., Lupyan, G., Boggio, P., Hickok, G., Gwilliams, L., Fernandino, L., Mirman, D., Chrysikou, E.G., Sandberg, C.W., Crutch, S.J., Pylkkaenen, L., Yee, E., Jackson, R.L., Rodd, J.M., Bedny, M., Connell, L., Kiefer, M., Kemmerer, D., de Zubicaray, G., Jefferies, E., Lynott, D., Siew, C.S.Q., Desai, R.H., Mcrae, K., Diaz, M.T., Bolognesi, M., Fedorenko, E., Kiran, S., Montefinese, M., Binder, J.R., Yap, M.J., Hartwigsen, G., Cantlon, J., Bi, Y. C., Hoffman, P., Garcea, F.E., Vinson, D., 2025. What we mean when we say semantic: toward a multidisciplinary semantic glossary. Psychon. Bull. Rev. 32, 243–280.

Renoult, L., Irish, M., Moscovitch, M., Rugg, M.D., 2019. From knowing to remembering: the Semantic-Episodic distinction. Trends Cogn. Sci. 23, 1041–1057.

Richards, B.A., Frankland, P.W., 2017. The persistence and transience of memory. Neuron 94, 1071–1084.

Rindal, E.J., DeFranco, R.M., Rich, P.R., Zaragoza, M.S., 2016. Does reactivating a witnessed memory increase its susceptibility to impairment by subsequent misinformation? J. Exp. Psychol. Learn. Mem. Cogn. 42, 1544–1558.

Ritchey, M., Libby, L.A., Ranganath, C., 2015. Cortico-hippocampal systems involved in memory and cognition: the PMAT framework. Prog. Brain Res. 219, 45–64.

Robertson, E.M., 2012. New insights in human memory interference and consolidation. Curr. Biol. CB 22. R66–71.

Robins, S.K., 2022. Episodic memory is not for the future. In: Sant'Anna, A., McCarroll, C., Michaelian, K. (Eds.), Current Controversies in the Philosophy of Memory. Routledge, pp. 166–184.

Robins, S.K., 2025. A place for the memory trace. In: Nadel, L., Aronowitz, S. (Eds.), Space, Time, and Memory. Oxford university Press.

Roediger, H.I.I.I., Abel, M., 2022. The double-edged sword of memory retrieval. Nat. Rev. Psychol. 1, 708–720.

Roy, D.S., Park, Y.G., Kim, M.E., Zhang, Y., Ogawa, S.K., DiNapoli, N., Gu, X.Y., Cho, J. H., Choi, H., Kamentsky, L., Martin, J., Mosto, O., Aida, T., Chung, K., Tonegawa, S., 2022. Brain-wide mapping reveals that engrams for a single memory are distributed across multiple brain regions. Nat. Commun. 13.

Rubin, D.C., Schrauf, R.W., Greenberg, D.L., 2003. Belief and recollection of autobiographical memories. Mem. Cogn. 31, 887–901.

Rugg, M.D., 2024. Frontoparietal contributions to retrieval. In: Kahana, M.J., Wagner, A. D. (Eds.), Handbook on Human Memory. Oxford University Press.

Rugg, M.D., Johnson, J.D., Uncapher, M.R., 2015. Encoding and retrieval in episodic memory: insights from fMRI. In: Addis, D.R., Barense, M., Duarte, A. (Eds.), The Wiley Handbook on the Cognitive Neuroscience of Memory. John Wiley & Sons, Ltd.

Rugg, M.D., Johnson, J.D., Park, H., Uncapher, M.R., 2008. Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. Prog. Brain Res. 169, 339–352.

Rugg, M.D., Srokova, S., 2024. Effects of age on neural reinstatement during memory retrieval, in, 2nd edition ed. In: Grafman, J.H. (Ed.), Module in Neuroscience and Biobehavioral Psychology, Encyclopedia of the Human Brain. Elsevier.

Rugg, M.D., Thompson-Schill, S.L., 2013. Moving forward with fMRI data. Perspect. Psychol. Sci. J. Assoc. Psychol. Sci. 8, 84–87.

Rugg, M.D., Vilberg, K.L., 2013. Brain networks underlying episodic memory retrieval. Curr. Opin. Neurobiol. 23, 255–260.

Rule, M.E., O'Leary, T., Harvey, C.D., 2019. Causes and consequences of representational drift. Curr. Opin. Neurobiol. 58, 141–147.

Sacripante, R., Logie, R.H., Baddeley, A., Della Sala, S., 2023. Forgetting rates of gist and peripheral episodic details in prose recall. Mem. Cogn. 51, 71–86.

Schacter, D.L., Addis, D.R., Buckner, R.L., 2007. Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. 8, 657–661.

Schacter, D.L., Chiu, C.Y.P., Ochsner, K.N., 1993. Implicit memory - a selective review. Annu. Rev. Neurosci. 16, 159–182.

Schacter, D.L., Thakral, P.P., 2024. Constructive memory and conscious experience. J. Cogn. Neurosci. 36, 1567–1577.

Schlichting, M.L., Preston, A.R., 2015. Memory integration: neural mechanisms and implications for behavior. Curr. Opin. Behav. Sci. 1, 1–8.

Scully, I.D., Napper, L.E., Hupbach, A., 2017. Does reactivation trigger episodic memory change? A meta-analysis. Neurobiol. Learn Mem. 142, 99–107.

Sederberg, P.B., Gershman, S.J., Polyn, S.M., Norman, K.A., 2011. Human memory reconsolidation can be explained using the temporal context model. Psychon. Bull. Rev. 18, 455–468.

Sekeres, M.J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., Moscovitch, M., 2016. Recovering and preventing loss of detailed memory: differential rates of forgetting for detail types in episodic memory. Learn. Mem. 23, 72–82.

Sekeres, M.J., Winocur, G., Moscovitch, M., 2018. The hippocampus and related neocortical structures in memory transformation. Neurosci. Lett. 680, 39–53.

Semon, R., 1904. Die mneme als erhaltendes prinzip im wechsel des organischen geschehens. Wilhelm Engelmann, London.

Shao, X.H., Li, A., Chen, C.S., Loftus, E.F., Zhu, B., 2023. Cross-stage neural pattern similarity in the hippocampus predicts false memory derived from post-event inaccurate information. Nat. commun. 14.

Shapson-Coe, A., Januszewski, M., Berger, D.R., Pope, A., Wu, Y.L., Blakely, T., Schalek, R.L., Li, P.H., Wang, S.H., Maitin-Shepard, J., Karlupia, N., Dorkenwald, S., Sjostedt, E., Leavitt, L., Lee, D., Troidl, J., Collman, F., Bailey, L., Fitzmaurice, A., Kar, R., Field, B., Wu, H., Wagner-Carena, J., Aley, D., Lau, J., Lin, Z.D., Wei, D.L., Pfister, H., Peleg, A., Jain, V., Lichtman, J.W., 2024. A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. Science 384.

Sharon, T., Moscovitch, M., Gilboa, A., 2011. Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. Proc. Natl. Acad. Sci. USA 108, 1146–1151.

Simmons, W.K., Ramjee, V., Beauchamp, M.S., McRae, K., Martin, A., Barsalou, L.W., 2007. A common neural substrate for perceiving and knowing about color. Neuropsychologia 45, 2802–2810.

Squire, L.R., Genzel, L., Wixted, J.T., Morris, R.G., 2015. Memory consolidation. Csh Perspect. Biol. 7.

Srokova, S., Hill, P.F., Rugg, M.D., 2022. The Retrieval-related anterior shift is moderated by age and correlates with memory performance. J. Neurosci. 42, 1765–1776.

Srokova, S., Hill, P.F., Elward, R.L., Rugg, M.D., 2021. Effects of age on goal-dependent modulation of episodic memory retrieval. Neurobiol. Aging 102, 73–88.

St Jacques, P.L., Montgomery, D., Schacter, D.L., 2015. Modifying memory for a museum tour in older adults: Reactivation-related updating that enhances and distorts memory is reduced in ageing. Memory 23, 876–887.

Steel, A., Billings, M.M., Silson, E.H., Robertson, C.E., 2021. A network linking scene perception and spatial memory systems in posterior cerebral cortex. Nat. Commun. 12.

Storm, B.C., Bjork, E.L., Bjork, R.A., 2012. On the durability of retrieval-induced forgetting. J. Cogn. Psychol. 24, 617–629.

Taylor, F.K., 1965. Cryptomnesia and plagiarism. Br. J. Psychiatry 111, 1111.

Teyler, T.J., Discenna, P., 1985. The role of hippocampus in memory - a hypothesis. Neurosci. Biobehav. Rev. 9, 377–389.

Tompary, A., Murty, V.P., 2025. Mechanisms of systems consolidation and their behavioral consequence, in: rugg, M.D. (Ed.), Systems neuroscience. Volume 3 of, 3rd edition ed Learning and Memory, A Comprehensive Reference. Elsevier.

Tonegawa, S., Morrissey, M.D., Kitamura, T., 2018. The role of engram cells in the systems consolidation of memory. Nat. Rev. Neurosci. 19, 485–498.

Tulving, E., 1972. Episodic and semantic memory. Organization of Memory. Academic Press, New York, pp. 381–403.

Tulving, E., 1974. Cue-Dependent forgetting. Am. Sci. 62, 74–82.

Tulving, E., 1983. Elements of Episodic Memory. Oxford University Press, New York.

Tulving, E., Pearlstone, Z., 1966. Availability versus accessibility of information in memory for words. J. Verbal Learn. Verbal Behav. 5, 381.

Uncapher, M.R., Rugg, M.D., 2009. Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. J. Neurosci. 29, 8270–8279.

van Dam, W.O., van Dijk, M., Bekkering, H., Rueschemeyer, S.A., 2012. Flexibility in embodied lexical-semantic representations. Hum. Brain Mapp. 33, 2322–2333.

Vatansever, D., Smallwood, J., Jefferies, E., 2021. Varying demands for cognitive control reveals shared neural processes supporting semantic and episodic memory retrieval. Nat. Commun. 12.

Wahlheim, C.N., Zacks, J.M., 2025. Memory updating and the structure of event representations. Trends Cogn. Sci. 29, 380–392.

Wardell, V., Palombo, D.J., 2024. Stability and malleability of emotional autobiographical memories. Nat. Rev. Psychol. 3, 393–406.

Willems, R.M., Casasanto, D., 2011. Flexibility in embodied language understanding. Front. Psychol. 2.

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., Anderson, M.C., 2015. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. Nat. Neurosci. 18, 582.

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. J. Cogn. Neurosci. 27, 679–691.

Winocur, G., Moscovitch, M., 2011. Memory transformation and systems consolidation. J. Int Neuropsychol. Soc. 17, 766–780.

Wixted, J.T., 2004a. On common ground: jost's (1897) law of forgetting and ribot's (1881) law of retrograde amnesia. Psychol. Rev. 111, 864–879.

Wixted, J.T., 2004b. The psychology and neuroscience of forgetting. Annu. Rev. Psychol. 55, 235–269.

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, 97–101.

Yassa, M.A., Stark, C.E.L., 2011. Pattern separation in the hippocampus. Trends Neurosci. 34, 515–525.

Yonelinas, A.P., Ranganath, C., Ekstrom, A.D., Wiltgen, B.J., 2019. A contextual binding theory of episodic memory: systems consolidation reconsidered. Nat. Rev. Neurosci. 20, 364–375.

Zacks, J.M., 2020. Event perception and memory. Annu. Rev. Psychol. 71 (71), 165–191.

Zeng, T.M., Tompary, A., Schapiro, A.C., Thompson-Schill, S.L., 2021. Tracking the relation between gist and item memory over the course of long-term memory consolidation. eLife 10.