PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Andonovski N, Sutton J, McCarroll CJ. 2024 Eliminating episodic memory? *Phil. Trans. R. Soc. B* **379**: 20230413.
https://doi.org/10.1098/rstb.2023.0413

Received: 5 February 2024
Accepted: 8 April 2024

One contribution of 22 to a theme issue 'Elements of episodic memory: lessons from 40 years of research'.

Subject Areas:
cognition

Keywords:
episodic memory, memory systems, eliminativism

Authors for correspondence:
John Sutton
e-mail: john.sutton@mq.edu.au
Nikola Andonovski
e-mail: ni.andonovski@gmail.com

THE ROYAL SOCIETY
PUBLISHING

Eliminating episodic memory?

Nikola Andonovski[1], John Sutton[2,3] and Christopher Jude McCarroll[4]

[1]Centre for Philosophy of Memory, IPhiG, Université Grenoble Alpes, Saint-Martin-d'Heres 38400, France
[2]Philosophy, Macquarie University, Sydney, New South Wales, Australia
[3]Philosophy, University of Stirling, Stirling, UK
[4]Institute of Philosophy of Mind and Cognition, National Yang Ming Chiao Tung University, Taipei, Taiwan

NA, 0000-0001-6957-0261; JS, 0000-0003-3046-9785; CJMC, 0000-0002-7555-1355

In Tulving's initial characterization, episodic memory was one of multiple memory systems. It was postulated, in pursuit of explanatory depth, as displaying proprietary operations, representations and substrates such as to explain a range of cognitive, behavioural and experiential phenomena. Yet the subsequent development of this research programme has, paradoxically, introduced surprising doubts about the nature, and indeed existence, of episodic memory. On dominant versions of the 'common system' view, on which a single simulation system underlies both remembering and imagining, there are no processes unique to memory to support robust generalizations with inductive potential. Eliminativism about episodic memory seems to follow from the claim that it has no dedicated neurocognitive system of its own. After identifying this under-noticed threat, we push back against modern eliminativists by surveying recent evidence that still indicates specialized mechanisms, computations and representations that are distinctly mnemic in character. We argue that contemporary realists about episodic memory can retain lessons of the common system approach while resisting the further move to eliminativism.

This article is part of the theme issue 'Elements of episodic memory: lessons from 40 years of research'.

1. Introduction

'Most, if not all, of our currently held ideas and theories about mental processes are wrong', Tulving admitted in his 1984 APA award address, and 'sooner or later in the future they will be replaced with more adequate concepts, concepts that fit nature better' [1, p. 386]. So, it was with modesty as well as ambition that, in *Elements of episodic memory* [2], he sought to bring that future closer, characterizing episodic memory as a specialized neurocognitive system underlying the capacity for remembering the personal past. It is the possession of this system that allows humans—indeed, only humans—to mentally travel back into the past and re-experience it in that warm and intimate way William James had described a century before.

Four decades after *Elements*, the sheer ubiquity of the concept of episodic memory may seem a reliable indicator of its good fit with nature. Textbooks and review articles routinely characterize episodic memory as a core mental capacity. The most notable development in the field, itself spearheaded by Tulving's pioneering work, has been the discovery of a close processing connection between remembering and imagining, often taken to indicate that a single neurocognitive system underlies both. Yet the deceptively simple shift to this 'common system' view introduces surprising doubts about the nature, and indeed existence, of episodic memory. We ask whether the new theories imply that the concept of episodic memory does not, after all, 'fit nature' well. Is the natural scientific development of Tulving's concept paradoxically leading to its elimination?

After summarizing the main features of Tulving's [2] original theory of episodic memory, we argue that on some interpretations this new generation of theories, positing a common system for remembering and imagination, lead to eliminativism: on such theories, episodic memory may not be a real kind. The consequences of this under-noticed implication may be severe. Eliminating episodic memory would potentially have dramatic costs, both practical and scientific, disrupting both our normative reliance on memory in social, legal and emotional contexts, and our empirical progress towards theories aiming at unification and consilience [3–5]. Given that the stakes are high, we should proceed with caution when faced with eliminativism. Retaining episodic memory may facilitate the integration of research on engrams, systems, learning, phenomenology and perhaps even autonoetic consciousness.

Having identified the threat, we go on to synthesize recent evidence, from various research programmes in the memory sciences, that challenges the eliminativist conclusion. This evidence, we argue, should make us reasonably optimistic about the existence of specialized episodic memory processes, potentially vindicating the core idea of *Elements*. We conclude by sketching options for revised forms of realism about episodic memory that still take on board the recent developments.

2. Tulving's *Elements*: episodic memory as an explanans

Tulving developed his theory of episodic memory, first systematically presented in *Elements*, within the multiple memory systems approach, which had emerged as a dominant research framework by the end of the 1970s.[1] On this approach, memory is not a unitary mental faculty but is composed of a number of memory systems: functionally distinct neurocognitive structures, constituted by sets of processes with proprietary principles of operation, representational kinds and neural substrates [7,8]. Memory systems constitute the basic *kinds* of memory and underpin the abilities of organisms to acquire, retain, retrieve and use information in a variety of productive ways. Episodic memory, Tulving argued, is a dedicated memory system that underpins the ability of individuals to remember events from their personal past and to re-experience them in a quasi-perceptual way, 'travel[ing] back into the past in their own minds' [2, p. 1]. Episodic memory is functionally distinct from semantic memory—a system underlying an individual's general or impersonal knowledge of the world—as well as from a number of procedural memory systems supporting the acquisition of various motor, perceptual and cognitive skills.

From the perspective of philosophy of science [9], the pursuit of explanatory depth lies at the heart of the multiple systems approach. The systems theorist aims to look beyond 'surface' variety to offer explanatory accounts of the underlying neurocognitive structures whose workings produce introspectively and behaviourally identifiable memory phenomena. In doing so, they aim to articulate principles that unify a variety of empirical generalizations, often from very different experimental paradigms. Such principles, along with the computational and representational properties of memory systems, account for the phenomena characterized by the available behavioural and phenomenological data. Memory systems, to employ a familiar philosophical idiom, are *natural* kinds [10,11]. They are causally organized clusters of properties of encoding, consolidation, storage and retrieval.[2] By design, the instantiation of some of these properties tends to cause the instantiation of the others, resulting in a robust and systematic correlation between them. By virtue of being natural kinds, memory systems afford productive theorizing, with natural kind terms like 'episodic memory' featuring in a wide range of stable empirical generalizations. They are thus to be distinguished from merely 'descriptive forms of memory'—such as verbal memory or recognition memory—which help describe and organize empirical facts but which do not carry such inductive potential [7, pp. 11–12].

Hypothesized to operate in a proprietary way, memory systems can be theorized about independently—in idealized isolation from their interaction with other systems. Yet, in reality, systems like episodic and semantic memory 'are closely interdependent and interact with one another virtually all the time' [2 p. 65]. Such interactions, which can be cooperative or competitive, occur in the production of most individual memories and are manifested in the performance of even simple experimental tasks. The relation between memory systems and memory tasks is many-to-many [2,7,12].

In *Elements*, episodic memory was thus posited as an explanans (that which does the explaining), its distinctive properties intended to account for core aspects of relevant memory phenomena. The theory sought to unify explanations of two main classes of such phenomena. The first is the ability of subjects to retain information about the properties of, and spatio-temporal relations between, items acquired on singular learning occasions, as manifested in (for example) classic verbal learning experiments [13]. The second, more prominently, concerned the phenomenology of remembering, which Tulving, following tradition, took to be marked by a 'definite affective tone that is uniquely and unmistakably one of the salient attributes of recollective experiences' [2, p. 48]. When remembering events from the personal past, subjects are immediately aware that they are re-experiencing events they have experienced before, a kind of awareness Tulving [14] would label autonoetic (i.e. self-knowing).[3] As a result, they are disposed to assert epistemic authority with respect to them, claiming knowledge of what happened even when unsure about the events' precise spatio-temporal location. This 'subjective veridicality', characteristic of personal recollection, is largely absent when we recall impersonal facts [2, p. 40].

We highlight three properties of the episodic memory system that Tulving introduced to account for these phenomena. First, the system was hypothesized to possess a proprietary store, in which information—paradigmatically, about the perceptible properties of experienced events—is registered directly [2, pp. 41–42]. Second, unlike semantic memory, episodic memory was taken to be relatively limited in its capacity to generate novel information not already present in the input [2, pp. 43–44]. Third, episodic retrieval was seen as a synergistic process, necessarily involving 'mixing' of information in the store, i.e. trace information, and information in the retrieval environment [2, pp. 171–176].[4] When the system functions properly, the retention of information encoded in the store affords a kind of 'immediate, or first-hand knowledge' of previously experienced events [2, p. 41], allowing subjects to successfully perform a variety of tasks that require such retention. Moreover, since the system is limited in its capacity to generate novel information, the very presence of some information in the store (at the point of retrieval) indicates that it was acquired on an epistemically relevant prior occasion. This partially grounds subjects' autonoetic awareness of a remembered event. Indeed, Tulving [14] argued, such awareness varies with the nature of the 'mix' between trace and cue information at retrieval: the larger the proportion of trace information, the greater the degree of autonoetic awareness.

In the pursuit of explanatory depth, *Elements* posited a specialized episodic memory system, whose functioning involved a delicate interaction between processes of information storage and dynamic, context-sensitive retrieval. The theory characterized episodic memory as a neurocognitive kind, supporting productive theorizing and causal generalizations with inductive potential. The theory, however, was not a taxonomic proposal for the classification of individual memor*ies*. Most memories emerge via the systematic interaction of multiple memory systems, a fact that led Tulving to characterize the task of unambiguously classifying a particular memory as episodic or semantic as 'uninteresting and lead[ing] nowhere' [15, p. 5]. The systems theorist aims to account for essential elements of memory phenomena—some of which, like the nature of conscious retrieval, may reflect underlying systemic differences—yet is fully aware of their behavioural and phenomenal diversity.[5]

3. Tulvingian eliminativism: from episodic memory to mental time travel but *not* back

Forty years after *Elements*, episodic memory has remained a key notion in the sciences of memory. Tulvingian themes dominate the contemporary discussion, as evidenced by the steady and well-documented rise in the number of articles dedicated to the examination of episodic memory, autonoetic consciousness and mental time travel [16–18]. A number of empirical developments have nevertheless gradually driven a notable and significant shift in focus.

Unsurprisingly, Tulving played a key role in initiating this new research. In 1985 [14], he reported the case of K.C., who had suffered a severe head injury in a traffic accident, resulting in widespread brain damage that included large bilateral hippocampal lesions. As a result, K.C. displayed a profound 'episodic' amnesia, being unable to recollect any personally experienced events from his past.[6] Intriguingly, he was also incapable of *imagining* possible future events, which suggested a close processing link between episodic memory and imagination [14]. The link was confirmed by subsequent neuropsychological studies reporting deficits exhibited by hippocampal amnesiacs in the imagination of novel events, whether located in the future or in the possible past [20–22]. In a parallel development, a slew of neuroimaging studies revealed that both remembering and imagination engage the default mode network (DMN), comprising the medial and lateral temporal lobes (hippocampus included), the medial prefrontal cortex, the posterior cingulate cortex and the inferior parietal lobule [23–25].[7] Psychological research also revealed developmental parallels between remembering and imagining [27] as well as analogous effects of temporal proximity and mental imagery abilities [28,29].

A new generation of theories, positing a common system underlying remembering and imagination—and, potentially, a number of other capacities—have emerged as a result. Tulving [30] again led the charge himself, characterizing episodic memory as a neurocognitive system for 'mental time travel' through subjective time. The operations of the system, he argued, allow subjects to re-experience the personal past in memory and pre-experience the personal future in imagination; both instances of autonoetically flavoured mental time travel (cf. [31]). Rival theories share the commitment to a common system, yet provide alternative accounts of its core operations. Schacter & Addis [32] characterize the system as one for 'constructive simulation' of possible events—in the past *and* future—a process that typically involves recombination of informational details acquired in the subject's previous experience (see also [26,33]). In the same spirit, Hassabis & Maguire [34] see remembering and imagination as underlaid by a system for the generation and maintenance of spatially coherent 'scene' representations (compare also [35,36]).

Importantly, the new theories inherit the spirit and methodological commitments of the multiple memory systems approach. Indeed, the positing of a common system—for mental time travel, constructive simulation or scene construction—is driven by the pursuit of core computational principles that may underlie a variety of cognitive activities and unify different empirical generalizations. The system, hypothesized to function in accordance with such principles, produces the behaviourally and introspectively accessible properties of personal remembering as well as future-oriented and counterfactual imagination. Proponents of these theories are happy to, and indeed often do, talk of episodic memory. Up to this point, we see continuities between the earlier and newer approaches.

But in fact, a significant terminological shift has gradually emerged, albeit often unnoticed, opening the door to doubts about the very existence of the system Tulving had originally sought to characterize. 'Episodic memory' is now most frequently used to refer to a product-state of the relevant system and *not* to the system itself. To put the point somewhat provocatively, 'episodic memory' now typically refers to the explanandum (that which needs to be explained), not the explanans. Addis [26, p. 70] provides a representative example, when she suggests that 'instead of conceptualising imagination as relying on episodic memory, we should consider memory and imagination as but two products of a constructive simulation system' (see also [34,37,38], among many others). We might decide to brush this off. After all, as Tulving repeatedly emphasized throughout his career, 'memory' can be used to refer both to a neurocognitive system and to its typical product. The shift, however, obscures a more significant issue: the dramatic change in the nature of the explanans.

On the picture presented in *Elements*, the operations of the episodic memory system were of direct explanatory significance for core properties of the target mnemic phenomena. This directness, indeed, licensed a verdict on function: an information-preserving system with limited generative capability seemed tailor-made for providing first-hand knowledge of personally experienced events. As theories shifted to a common system underlying a variety of more disparate phenomena, the hypothesized operations became more inferentially distant from (descriptions of) these phenomena (see [9, §4]). While details vary, four important changes are worth highlighting. First, the new theories do not typically posit a proprietary 'episodic' store, dedicated to the retention of information acquired on singular, epistemically relevant occasions. Rather, the system is considered to utilize information from a variety of sources in a flexible and context-sensitive manner. Second, the system is seen as strongly generative, routinely and systematically producing information not present in the original input. Third, and relatedly, episodic retrieval is characterized as a more thoroughly reconstructive process, involving the formation of distinct kinds of representations, carefully calibrated for a variety of mnemic and imaginative uses. The process, typically seen as simulational in nature, is taken to constitutively depend on the operations of the DMN. Given that the DMN seems to be involved in a range of episodic simulations, this structure is understood as 'the brain's simulation system' [26, p. 71], and episodic remembering is taken to be one particular operation of this system.[8] Finally, there has been a significant shift in the characterization of the system's proper function, now systematically linked to the enhancement of subjects' predictive prowess and ability to imagine the future. The episodic system, in short, is not really *for* remembering the personal past (for discussions, see [26,30–35,38–41]).

Crucially, the mental time travel/simulation system is hypothesized to function uniformly, with the same set of neurocognitive operations underlying personal remembering and different forms of imagination. This *identity-of-process* hypothesis, widely accepted by systems theorists, is nicely articulated by Addis ([26, p. 70], emphases added):[9]

> From a neurocognitive perspective, it is likely that an event representation is physically instantiated in the brain in *exactly the same way irrespective of whether it is remembered or imagined:* it is a set of connections between the nodes representing perceptual content, semantic information and related schemas... [In remembering and imagining events] *the same simulation process is engaged and the same types of information are drawn upon*.

Hence, not only does a common neurocognitive system underlie both remembering and imagining; further, there are no component processes of the system dedicated solely to the production of memories or imaginings. Both involve the same strongly generative process of reconstructive retrieval in which information from different sources is pieced together in an event representation that aims to satisfy a number of constraints (such as relevance, accuracy, general plausibility, consistency with knowledge of past/future events). The identity-of-process hypothesis does not preclude the existence of *some* differences between memories and imaginings, likely resulting from the task-sensitive differential engagement of the constituent operations of the system. These, however, are differences in degree and not in kind [25,26,33,38,39].

These developments turn the new systems theories decisively in the direction of eliminativism about episodic memory. Tulving's original ambition of accounting for autonoetically flavoured remembering by appeal to a specialized *memory* system seems to have been given up [9]. At least prima facie, a neurocognitive system for mental time travel or constructive simulation—lacking a dedicated store and selected for its effects on subjects' ability to imagine the future—is not such a system. Moreover, given identity-of-process, there is no specialized component process of this system uniquely engaged during personal remembering. The consequences are far-reaching. The multiple systems approach had treated memory systems and their constituent processes as constituting the basic memory kinds, which afford productive and inductively fruitful theorizing. But if, in contrast, the identity-of-process hypothesis held by common system theorists is correct, then episodic memory is *not* a real kind; indeed, it is not even a sub-kind of mental time travel or constructive episodic simulation. As a result, there are no causal generalizations about episodic memory that carry inductive potential.

This striking conclusion has indeed surfaced in recent philosophical discussions: in particular, Michaelian's [38] variant of the simulation theory has been diagnosed as eliminativist [44,45]. Such eliminativism about episodic memory does not entail that humans do not engage in remembering the personal past, a mental activity that may exhibit a number of unique characteristics. It does, however, entail that there is no dedicated neurocognitive episodic memory system that underlies remembering. In this regard, remembering may be less like visual perception and more like cherishing, overlooking or disregarding.[10]

We should be careful here. An eliminativist's modus ponens is a revisionist's modus tollens. A revisionist, looking at the new developments, may insist on identifying episodic memory with the 'broader' neurocognitive system. Indeed, this is the route taken by Tulving [30, p. 9] who argued that episodic memory just *is* a system that 'makes possible mental time travel through subjective time—past, present, and future'.[11] This route is certainly open. As generations of philosophers have learned from Quine, an empirical discovery can be redescribed as a change in meaning. Our goal here is thus not to legislate the use of 'episodic memory' in the sciences of memory. It is rather to highlight the significant changes in theorizing about the system hypothesized to support personal remembering. These involve a shift from an information-preserving system that affords first-hand knowledge of the personal past to a more general system for event representation whose evolutionary function may indeed not be mnemic at all. We take the characteristic reluctance of theorists to use 'episodic memory' to refer to such a system as a reliable indicator that this shift is indeed quite significant. A revisionist who takes this lesson on board is, for our purposes at least, not essentially different from an eliminativist about episodic memory. To be clear, we do *not* think that the common system hypothesis entails eliminativism. In fact, we take it as plausible that some forms of common system theory will retain some form of realism about episodic memory, thus resisting eliminativism: there are a range of less radically revisionary ways to incorporate the key lessons of common system theories. But our argument here is, nevertheless, that the currently predominant versions of common system theory do in fact eliminate episodic memory as a real kind.[12]

A related issue pertains to the relationship between the underlying system and the target mnemic phenomena. The increase in inferential distance between the two, of course, does not entail that the hypothesized system is not causally involved in producing states of personal, autonoetic remembering. What has become increasingly evident, however, is that such production likely involves more widespread interactions with other cognitive and meta-cognitive processes, interactions that may be context-sensitive and not easily accounted for by the core explanatory principles of the new system theories. This realization motivates Mahr's [43,47] recent attempt to account for the ways in which a general episodic simulation system produces the characteristically past-oriented, specific and self-involving memories of the personal past. Mahr reserves a key role for metacognitive monitoring mechanisms, which regularly interact with the outputs of the simulation system and whose function—particularly in attributions of 'mnemicity'—is calibrated via an extended process of social and cultural learning [48].

The new common system theories are often presented as a natural continuation of Tulving's work. But, we have argued, these 'neo-Tulvingian' theories have paradoxically come to countenance or embrace an eliminativism about episodic memory. Perhaps this should not come as a surprise. If scientific development is regulated by an ideal of theoretical unification, and is not beholden to common-sense modes of explanation, then we should expect the occasional elimination, or at least radical revision, of familiar categories [49]. Yet in this case it may be too early to admit defeat. Other recent developments in the memory sciences may indeed vindicate core ideas of *Elements*.

4. Back to episodic memory?

The sciences of memory constitute a broad and increasingly interdisciplinary field, housing a variety of research programmes and traditions. Empirical developments may thus motivate different responses to the common system and identity-of-process hypotheses. One critical response would involve highlighting phenomenological and behavioural differences between the target phenomena, personal remembering and imagination. Since common system theorists can allow for such differences, however, such evidence is unlikely to be decisive. What seems necessary to challenge the eliminativist verdict is evidence for dedicated mechanisms for episodic memory storage, specialized computational operations or representational kinds as well as relevant differences in underlying neural substrates. Theoretical unification under the common system and identity-of-process hypotheses can be accomplished only if there are no robust generalizations with inductive potential that pertain to a (sub)set of processes that are, in some recognizable way, distinctly mnemic in character. Even with that in mind, the sheer diversity of theoretical and empirical approaches is likely to afford different strategies against eliminativism about episodic memory. In this section, we briefly address just *two* such strategies, one negative and one positive, developing the latter more fully.

First, a key *negative* task for the critic would be to undermine confidence in the existence of a common, functionally specialized process underlying both remembering and imagination. In a recent article [50], De Brigard targets the putative role of the DMN in constructive simulation and scene construction theories of the common system. He presents evidence for the engagement of the DMN in processes that do not involve simulation of events or scenes, such as interoception [51] or semantic processing [52].[13] Moreover, he points to the existence of spatial mappings—hypothesized to anchor event simulations—outside this network (for example [53]), to the ability of individuals with hippocampal damage to generate spatially coherent representations [54], and to the differential engagement of the hippocampus in different forms of event representations, indicating likely processing differences.[14] This last consideration interplays well with recent neural and behavioural evidence of such differences (see below) as well as with neuroimaging data showing that, despite substantial overlap between the recollective and 'general semantic' networks, the hippocampus plays a hub role in the former but does not feature in the latter [16].

The *positive* task, to which we now turn, involves proposing inductively promising generalizations about mnemic processes. A key component of Tulving's [2] account was the retention of information acquired on singular learning occasions. Recent technological advances have afforded a more direct investigation of the discrete neural vehicles underlying such retention. Molecular and transgenic 'engram technology' has allowed researchers to isolate neural populations coding for specific properties of learning stimuli and to demonstrate their causal relevance for subsequent retrieval, results that have reinvigorated the commitment to discrete memory traces with unique—and empirically tractable—causal histories [55,56]. Indeed, experimenters have begun to trace the (changing) neural organization of traces, showing relative diachronic stability in certain areas of the hippocampus, paired with greater cortical engagement over time [57].[15] Engram-technological manipulations are most commonly employed in rodent conditioning studies, but there are ongoing attempts to design experimental tasks that require (more) complex representations and computational operations as well as to establish the relevance of the findings to declarative memory (for critical discussion, see [58]). In humans, the employment of representational similarity analysis and deep neural network techniques has provided insights into the structure, format and informational profiles of neural event representations [59,60]. The results show re-expression of content-sensitive patterns of neural activity at retrieval, alongside systematic transformations from perceptual to conceptual formats, raising important questions about the relation between the two processes (cf. [61]).

The studies have implicated a variety of mechanisms likely to play a role in information retention and transformation. The hippocampus, unsurprisingly, has been at the centre of attention. Different hippocampal cell types have been shown to be responsive to specific stimulus features and are believed to play important roles in binding informational content into diachronically stable event representations (for a review, see [62]). Relatedly, single neurons in the hippocampus have also been shown to respond to perceptual event boundaries, with their activity predictive of subsequent memory performance [63]. Moreover, there is some evidence for the differential role of hippocampal subfields. Cell populations in the DG and CA3 subfields have been implicated in the encoding of distinct, pattern-separated representations, affording rapid 'episodic' learning in the absence of interference. CA1 populations, likely involved in such learning, have also been linked to more overlapping representations, supporting information integration and generalization across events [64–66].

Behavioural evidence complements this picture nicely. Studies show systematic retention of information not only about the sensory and spatio-temporal properties of experienced events—in a characteristically 'all-or-none' way [67]—but also about event boundaries, with perceptual segmentation shown to affect long-term memory organization [68]. Importantly, sleep-dependent consolidation, believed to support generalization, also appears to benefit retention of information about specific events. Hence, in a memory similarity task, Hanert *et al.* [69] found that sleep in humans stabilizes pattern separation performance. Schapiro *et al.* [70] similarly reported retention of exemplar-specific information during consolidation, paralleled by improvements in memory for semantic category structure. Results of this kind suggest that discrete traces, carrying information about specific events or items, are likely often preserved in consolidation, despite documented processes of trace transformation and schematization. As Heinen *et al*. [60] point out, consolidation may involve *proliferation* of memory representations with different formats and accessibility conditions, not literal transformation of specific traces from one format to another (cf. [71]). The balance between preservation and reconstruction described in *Elements* may not be a thing of the past, after all.

Theoretical and empirical treatments of memory traces have typically paid little attention to the specific systems, and component processes, involved in information storage. This is unfortunate, as a systems perspective is of key importance: it affords the specification of the computational problems trace-manipulating systems are designed to solve as well as of their functional and resource constraints (cf. [72]). Recently, theorists have attempted to remedy this problem, with a variety of proposals about the roles discrete event representations play in cognition. There has been a characteristic emphasis on one-shot learning—i.e. 'episodic control'—a process that involves the use of information about specific events, including actions performed, for learning an appropriate action sequence [73]. Modelling work, building on developments in deep reinforcement learning, has illustrated the computational advantages of employing episodic control processes, hypothesized to rely on dedicated hippocampal mechanisms [74,75]. In a notable recent study, Zeng *et al*. [76] provided evidence that episodic control benefits spatial learning, compared with replay-driven processes, particularly when the task is sufficiently complex and the number of learning trials is limited.

The epistemic benefits of using pattern-separated representations of individual events have also been examined in the context of associative learning, with the retention of specific details hypothesized to afford flexible use of information at retrieval [77]. Zhou *et al.* [78] examined the integration of related experiences in humans, using a hippocampally dependent associative inference task. Their results indicated that subjects likely employ both pattern-separated and integrated event representations, with characteristic advantages in different retrieval contexts. Specifically, the retention of pattern-separated representations was found to be beneficial in circumstances that do not require rapid recognition of item relatedness, for example in explicit inference. (In contrast, integration of experiences via interleaved learning gave rise to fast, automatic recognition.) Comparing computational models of memory integration, Zhou *et al*. concluded that only models that include both kinds of representations can account for the available behavioural data.

The evidence presented above motivates a careful reassessment of the identity-of-process view. A tentative hypothesis posits the existence of a set of episodic *memory* processes—possibly constituting a memory system—dedicated to the preservation, maintenance and retrieval of discrete event representations. Tightly integrated, the processes are relatively limited in their ability to generate novel information, yet afford flexible and context-sensitive recall. The preservation of discrete memory traces is constitutively involved in remembering the personal past, with information from first-hand experiences systematically prioritized at retrieval. Such preservation, to echo Tulving & Watkins, 'is a necessary condition for the subsequent retrieval of information about the event' [79, p. 261]. Despite the substantial similarities—plausibly due to the engagement of simulative processes of event representation—discrete memory traces are likely *not* constitutively involved in future-oriented or counter-factual imagination. While these often rely on individual traces, their production is not 'diachronically' constrained in the same way ([44]; cf. [80]). Additional support for this hypothesis may come from more recent studies comparing the neural underpinnings and development of remembering and episodic future thought. Building on prior work, Østby *et al*. [81] found that the functional connectivity of the DMN, in adolescents and young adults, was strongly related to the subjective quality of personal remembering, yet only marginally related to episodic future thought. This finding fits well with the growing number of studies reporting significant differences in the developmental trajectories of the two capacities (for a review, see [82]).

Importantly, positing a specialized set of processes for the preservation of event representations does not commit us to an 'archival' or non-constructivist view of episodic memory. As we have tried to emphasize throughout the paper, memories emerge via the systematic interaction of a variety of processes. While preservative processes may be limited in their capacity to generate novel information, such generation occurs regularly via synergistic trace–cue interactions. Moreover, event representations with different formats and informational profiles are likely systematically maintained and manipulated, from encoding to consolidation and retrieval. These representations may be stored independently, yet they contribute jointly to the informational structure of individual memories, with their dynamic interactions dependent on a number of factors, such as attention, time, task demands and prior knowledge [4,60,71,83].

While there are reasons for optimism about the place of episodic memory in science, we should proceed with caution. Process hypotheses owe answers to notoriously difficult questions about the individuation of processes and their integration in distinct neurocognitive systems. Individuating a computational process requires careful, and likely theory-laden, abstraction from the 'unfiltered' causal structure of the world [84], an endeavor that has proven difficult across the memory sciences [85]. Questions about integration, which had worried Tulving throughout his career, are arguably even more pressing. While functionally specialized, memory systems were hypothesized to have 'fuzzy boundaries, have overlapping constituent processes, and [to] interact with one another in intricate ways' [7, p. 18]. Establishing whether a set of processes are sufficiently integrated to constitute a system requires specification of the degree of integration necessary for constituency, with pragmatic factors likely to play a major role. Further, the integration of a set of processes in a common neurocognitive system might essentially depend on the experiences of the organism, thus constituting a kind of developmental achievement [86]. Getting reliable evidence for the actual form and degree of integration is difficult, particularly given the hypothesized intricate interactions between systems. While emerging neuroimaging and network data may alleviate some of these problems, establishing the proper interpretation and relevance of such data remains a considerable challenge (for a discussion of these issues, see [9]).[16]

A proper assessment of the rival views, and of their consilience with the ideas outlined in *Elements*, thus requires navigating complex empirical and conceptual terrains. Future work, aiming to advance the debate, would have to examine the architectural connections between 'core' episodic processes (such as information retention, constructive simulation and autonoesis), their possible integration in neurocognitive systems, and the principles that govern their interactions in the production of mnemic and imaginative states. Such work, the methodological underpinnings of which may already be in place [16,17,43], promises to be arduous and will plausibly involve a further increase in inferential distance between the relevant explanantia and explananda. Yet we have no other choice. Engaging in this work is the only way to hasten the arrival of a future in which we know whether Tulving's [2] signature idea of a specialized episodic memory system is wrong—and perhaps even whether it is right.

5. Conclusion

The notion of episodic memory has proven remarkably resilient. Somewhat surprisingly, however, this resilience may not be a reliable indicator of its good fit with nature. For Tulving [2], the notion stood for a hypothesized neurocognitive system underlying the familiar experience of remembering the personal past. In the intervening decades, empirical developments have necessitated significant amendments to this picture. Keeping the systems perspective, contemporary theories posit a common system for remembering and imagining, often retaining the notion of episodic memory to refer to individual memories produced by this system. In this paper, we examined this shift in perspective, diagnosing the new theories as eliminativist in tendency. On many common system views, we argued, episodic memory does not constitute a real kind, appeals to which may ground hypotheses with inductive potential. We then moved on to examine and systematize recent evidence that may challenge this conclusion. Such evidence comes from a number of research programmes and provides preliminary support for the existence of a set of processes, dedicated to the retention and maintenance of discrete representations of personally experienced events. Future work, which will have to negotiate difficult conceptual and empirical questions, will reveal how these processes are involved in remembering and imagination and whether they constitute a unified episodic memory system.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. This article has no additional data.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. N.A.: conceptualization, formal analysis, investigation, methodology, writing—original draft, writing—review and editing; J.S.: conceptualization, formal analysis, investigation, methodology, writing—review and editing; C.J.McC.: conceptualization, formal analysis, investigation, methodology, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. N.A.'s research was funded by the European Union's Horizon research and innovation programme under the Marie Sklodowska-Curie grant agreement 101062754. J.S.'s work was supported by a Marie Sklodowska-Curie fellowship at the Paris Institute for Advanced Study, by Templeton grant no. 61924 'Concepts as Dynamic Assemblages' awarded to Agustín Fuentes and Greg Downey, and by a Leverhulme visiting fellowship at the University of Stirling. C.McC.'s work was also supported by a Taiwan National Science and Technology Council grant (112-2410 H-A49-084-MY3), and he gratefully acknowledges this funding.

Acknowledgements. We thank the special issue editors for their support and assistance, and the two anonymous reviewers for extremely helpful comments. N.A. thanks José Camillo, Becko Copenhaver, Kirk Michaelian and James Openshaw for helpful discussions. C.McC. thanks Paloma Muñoz for her help and support.

Endnotes

[1]The notion of episodic memory first appeared in Tulving [6] in the context of the hypothesis that episodic and semantic memory are distinct information-processing systems with characteristic computational properties. At that point, the hypothesis was treated as a 'pretheoretical position whose major usefulness may turn out to lie in facilitating theory construction' [6, p. 384] and was not explicitly linked to a search for systems or natural psychological kinds.

[2]For the purposes of this paper, we do not want to commit to any particular conception of natural kinds. While we have sympathies for Boyd's [10] homeostatic property cluster view, we have some reservations about the necessity of a 'homeostatic' mechanism that keeps the relevant properties clustered. A loosening of Boyd's account, perhaps along the lines of Khalidi [11], seems to us a step in the right direction.

[3]The immediacy of awareness entails the absence of explicit inference, *not* immunity to error. Subjects undergoing autonoetic experiences can, of course, be mistaken about (having previously experienced) a past event.

[4]Technically, it is *ecphory*—a process of activation of a latent trace via conversion of trace and cue information—that is synergistic. Ecphory is only one element of retrieval [2, pp. 175–178]. So as not to complicate things too much, we gloss over this issue in the main text.

[5]Space prohibits a detailed investigation of the claims advanced in this section. For a fuller historical and philosophical treatment of the original theory in relation to the empirical research that it motivated and rested on, see [9].

[6]K.C.'s general world knowledge, including much of his knowledge about himself, was surprisingly preserved, which suggested that the amnesia was characteristically *episodic*. For more details, see Rosenbaum *et al*. [19].

[7]Remembering and imagining activated the DMN relative to a semantic control task, such as thinking about the definitions of cue words and generating words related to the cue. The DMN is thought to be involved in a number of different tasks including when individuals are not engaged in a cognitive activity that requires attention to external sources, or when individuals are mind-wandering [26]. We discuss the DMN further below.

[8]Addis [33, p. 234] notes that the DMN is not only involved in forms of mental time travel, but is also engaged in other forms of cognition including creativity, theory of mind, narrative comprehension and event perception. These diverse forms of cognitive activity are nonetheless united, she thinks, by involving simulations—the mental renderings of experience.

[9]See also, for example De Brigard [39, p. 173]; Addis [33, p. 239]; Schacter [42, p. 40]; Mahr [43, p. 3].

[10]We assume, of course, that none of these states are underlaid by dedicated brain systems.
[11]Michaelian [46] offers a comparably radical revision, arguing that episodic remembering is constituted by simulation of a past event, regardless of whether the event has been previously experienced.
[12]Our reviewers suggest that theorists, faced with these difficulties, may shift to a pragmatist or normative conception of episodic memory. We agree. Indeed, both descriptive and normative perspectives may be appropriate in different explanatory contexts [3,5]. Here, however, we examine the development of 'neo-Tulvingian' theories which purport to inherit the methodological commitments of Tulving's multiple memory systems approach and treat episodic memory as a *natural* kind.
[13]De Brigard raises further worries about understanding the DMN as the simulation system of the brain. He notes that many non-human animals possess a DMN but thinks it is highly unlikely they engage in complex episodic simulations. He also emphasizes that not all episodic constructive processes depend to the same extent on core regions of the DMN; for example, the DMN fails to be preferentially recruited in cognizing episodic counterfactual situations that involve objects as opposed to people, and it is also not recruited when thinking counterfactually about what an unfamiliar person in an unfamiliar situation would have done.
[14]Importantly, subjects with hippocampal damage exhibit substantial deficits in episodic *memory* tasks, despite their preserved ability to generate spatially coherent representations (see the discussion in [54]). These deficits are plausibly (partially) due to impairments of specialized preservative processes of the kind we examine below.
[15]Specifically, Refaeli *et al*. found that a core ensemble in the CA1 region of the dorsal hippocampus was *necessary* for retrieval of both recent and remote memories.
[16]On Tulving's conception, memory systems are constituted by tightly integrated constituent computational processes. Hence, positing a memory system requires evidence not only of specialized processes (which may or may not be proprietary to the system) but also of their systematic integration. As we have tried to emphasize, the latter kind of evidence may be particularly difficult to obtain. We thank a reviewer for prompting us to add this clarification.

References

1. Tulving E. 1985 How many memory systems are there? *Am. Psychol*. **40**, 385–398. (doi:10.1037//0003-066X.40.4.385)
2. Tulving E. 1983 *Elements of episodic memory*. Oxford, UK: Oxford University Press.
3. Craver CF. 2020 Remembering: epistemic and empirical. *Rev. Philos. Psychol*. **11**, 261–281. (doi:10.1007/s13164-020-00469-7)
4. Andonovski N. 2021 Memory as triage: facing up to the hard question of memory. *Rev. Philos. Psychol*. **12**, 227–256. (doi:10.1007/s13164-020-00514-5)
5. McCarroll CJ, Michaelian K, Nanay B. 2024 Explanatory contextualism about episodic memory: towards a diagnosis of the causalist-simulationist debate. *Erkenntnis* **89**, 2273–2301. (doi:10.1007/s10670-022-00629-4)
6. Tulving E. 1972 Episodic and semantic memory. In *Organization of memory* (eds E Tulving, W Donaldson), pp. 381–402. Oxford, UK: Academic Press.
7. Schacter DL, Tulving E. 1994 What are the memory systems of 1994? In *Memory systems 1994* (eds DL Schacter, E Tulving). Cambridge, MA: MIT Press. (doi:10.7551/mitpress/4545.001.0001)
8. Squire LR. 2004 Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem*. **82**, 171–177. (doi:10.1016/j.nlm.2004.06.005)
9. Andonovski N. 2023 Autonoesis and the Galilean science of memory: explanation, idealization, and the role of crucial data. *Eur. J. Philos. Sci*. **13**, 42. (doi:10.1007/s13194-023-00548-3)
10. Boyd R. 1991 Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philos. Stud*. **61**, 127–148. (doi:10.1007/BF00385837)
11. Khalidi MA. 2018 Natural kinds as nodes in causal networks. *Synthese* **195**, 1379–1396. (doi:10.1007/s11229-015-0841-y)
12. Tulving E. 1991 Concepts of human memory. In *Memory: organization and locus of change* (eds LR Squire, NM Weinberger, G Lynch, JL McGaugh). Oxford, UK: Oxford University Press. (doi:10.1093/oso/9780195069211.003.0001)
13. Tulving E, Madigan SA. 1970 Memory and verbal learning. *Annu. Rev. Psychol*. **21**, 437–484. (doi:10.1146/annurev.ps.21.020170.002253)
14. Tulving E. 1985 Memory and consciousness. *Can. Psychol. / Psychol. Can*. **26**, 1–12. (doi:10.1037/h0080017)
15. Tulving E. 2002 Episodic memory: from mind to brain. *Annu. Rev. Psychol*. **53**, 1–25. (doi:10.1146/annurev.psych.53.100901.135114)
16. Renoult L, Irish M, Moscovitch M, Rugg MD. 2019 From knowing to remembering: the semantic–episodic distinction. *Trends Cogn. Sci.* **23**, 1041–1057. (doi:10.1016/j.tics.2019.09.008)
17. Dafni-Merom A, Arzy S. 2020 The radiation of autonoetic consciousness in cognitive neuroscience: a functional neuroanatomy perspective. *Neuropsychologia* **143**, 107477. (doi:10.1016/j.neuropsychologia.2020.107477)
18. Miloyan B, McFarlane KA, Suddendorf T. 2019 Measuring mental time travel: is the hippocampus really critical for episodic memory and episodic foresight? *Cortex* **117**, 371–384. (doi:10.1016/j.cortex.2019.01.020)
19. Rosenbaum RS, Köhler S, Schacter DL, Moscovitch M, Westmacott R, Black SE, Gao F, Tulving E. 2005 The case of KC: contributions of a memory-impaired person to memory theory. *Neuropsychologia* **43**, 989–1021. (doi:10.1016/j.neuropsychologia.2004.10.007)
20. Klein SB, Loftus J, Kihlstrom JF. 2002 Memory and temporal experience: the effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Soc. Cogn*. **20**, 353–379. (doi:10.1521/soco.20.5.353.21125)
21. Hassabis D, Kumaran D, Vann SD, Maguire EA. 2007 Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl Acad. Sci.* **104**, 1726–1731. (doi:10.1073/pnas.0610561104)
22. Mullally SL, Maguire EA. 2014 Memory, imagination, and predicting the future: a common brain mechanism? *Neuroscientist* **20**, 220–234. (doi:10.1177/1073858413495091)
23. Okuda J *et al*. 2003 Thinking of the future and past: the roles of the frontal pole and the medial temporal lobes. *Neuroimage* **19**, 1369–1380. (doi:10.1016/s1053-8119(03)00179-4)
24. Addis DR, Moscovitch M, McAndrews MP. 2007 Consequences of hippocampal damage across the autobiographical memory network in left temporal lobe epilepsy. *Brain* **130**, 2327–2342. (doi:10.1093/brain/awm166)
25. Schacter DL, Addis DR, Hassabis D, Martin VC, Spreng RN, Szpunar KK. 2012 The future of memory: remembering, imagining, and the brain. *Neuron* **76**, 677–694. (doi:10.1016/j.neuron.2012.11.001)
26. Addis DR. 2018 Are episodic memories special? On the sameness of remembered and imagined event simulation. *J. R. Soc. New Zeal*. **48**, 64–88. (doi:10.1080/03036758.2018.1439071)
27. Atance CM, O'Neill DK. 2001 Episodic future thinking. *Trends Cogn. Sci.* **5**, 533–539. (doi:10.1016/S1364-6613(00)01804-0)

28. D'Argembeau A, Van der Linden M. 2004 Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: influence of valence and temporal distance. *Consc. Cogn.* **13**, 844–858. (doi:10.1016/j.concog.2004.07.007)
29. D'Argembeau A, Van der Linden M. 2006 Individual differences in the phenomenology of mental time travel: the effect of vivid visual imagery and emotion regulation strategies. *Conscious. Cogn*. **15**, 342–350. (doi:10.1016/j.concog.2005.09.001)
30. Tulving E. 2005 Episodic memory and autonoesis: uniquely human? In *The missing link in cognition: origins of self-reflective consciousness* (eds HS Terrace, J Metcalfe). Oxford, UK: Oxford University Press. (doi:10.1093/acprof:oso/9780195161564.003.0001)
31. Suddendorf T, Corballis MC. 2007 The evolution of foresight: what is mental time travel, and is it unique to humans? *Behav. Brain Sci*. **30**, 299–313; (doi:10.1017/S0140525X07001975)
32. Schacter DL, Addis DR. 2007 The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Phil. Trans. R. Soc. B* **362**, 773–786. (doi:10.1098/rstb.2007.2087)
33. Addis DR. 2020 Mental time travel? A neurocognitive model of event simulation. *Rev. Philos. Psychol*. **11**, 233–259. (doi:10.1007/s13164-020-00470-0)
34. Hassabis D, Maguire EA. 2009 The construction system of the brain. *Phil. Trans. R. Soc. B* **364**, 1263–1271. (doi:10.1098/rstb.2008.0296)
35. Rubin DC, Umanath S. 2015 Event memory: a theory of memory for laboratory, autobiographical, and fictional events. *Psychol. Rev*. **122**, 1–23. (doi:10.1037/a0037907)
36. Cheng S, Werning M, Suddendorf T. 2016 Dissociating memory traces and scenario construction in mental time travel. *Neurosci. Biobehav. Rev.* **60**, 82–89. (doi:10.1016/j.neubiorev.2015.11.011)
37. Conway MA. 2009 Episodic memories. *Neuropsychologia* **47**, 2305–2313. (doi:10.1016/j.neuropsychologia.2009.02.003)
38. Michaelian K. 2016 *Mental time travel: mental time travel: episodic memory and our knowledge of the personal past*. Cambridge, MA: MIT Press. (doi:10.7551/mitpress/10591.001.0001)
39. De Brigard F. 2014 Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese* **191**, 155–185. (doi:10.1007/s11229-013-0247-7)
40. Klein SB. 2016 Autonoetic consciousness: reconsidering the role of episodic memory in future-oriented self-projection. *Q. J. Exp. Psychol.* **69**, 381–401. (doi:10.1080/17470218.2015.1007150)
41. Schulz AW, Robins S. 2023 Episodic memory, simulated future planning, and their evolution. *Rev. Philos. Psychol.* **14**, 811–832. (doi:10.1007/s13164-021-00601-1)
42. Schacter DL. 2022 The seven sins of memory: an update. *Memory* **30**, 37–42.
43. Mahr JB. 2023 How to become a memory: the individual and collective aspects of mnemicity. *Top. Cogn. Sci*. **16**, 225–240. (doi:10.1111/tops.12646)
44. McCarroll CJ. 2020 Remembering the personal past: beyond the boundaries of imagination. *Front. Psychol*. **11**, 585352. (doi:10.3389/fpsyg.2020.585352)
45. Hoerl C. 2022 A knowledge-first approach to episodic memory. *Synthese* **200**, 376. (doi:10.1007/s11229-022-03702-1)
46. Michaelian K. 2024 Radicalizing simulationism: remembering as imagining the (nonpersonal) past. *Philos. Psychol*. **37**, 1170–1196. (doi:10.1080/09515089.2022.2082934)
47. Mahr JB. 2020 The dimensions of episodic simulation. *Cognition* **196**, 104085. (doi:10.1016/j.cognition.2019.104085)
48. Mahr JB, van Bergen P, Sutton J, Schacter DL, Heyes C. 2023 Mnemicity: a cognitive gadget? *Perspect. Psychol. Sci*. **18**, 1160–1177. (doi:10.1177/17456916221141352)
49. Collins J. 2007 Meta-scientific eliminativism: a reconsideration of Chomsky's review of skinner's verbal behavior. *Br. J. Philos. Sci*. **58**, 625–658. (doi:10.1093/bjps/axm041)
50. De Brigard F. Simulationism and memory traces. In *Memory, space and time* (eds S Aronowitz, L Nadel). Oxford, UK: Oxford University Press.
51. Kleckner IR, Zhang J, Touroutoglou A, Chanes L, Xia C, Simmons WK, Quigley KS, Dickerson BC, Barrett LF. 2017 Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nat. Hum. Behav*. **1**, 0069. (doi:10.1038/s41562-017-0069)
52. Lanzoni L, Ravasio D, Thompson H, Vatansever D, Margulies D, Smallwood J, Jefferies E. 2020 The role of default mode network in semantic cue integration. *Neuroimage* **219**, 117019. (doi:10.1016/j.neuroimage.2020.117019)
53. Long X, Zhang SJ. 2021 A novel somatosensory spatial navigation system outside the hippocampal formation. *Cell Res*. **31**, 649–663. (doi:10.1038/s41422-020-00448-8)
54. De Brigard F, Gessell B. 2016 Time is not of the essence: understanding the neural correlates of mental time travel. In *Seeing the future: theoretical perspectives on future-oriented mental time travel* (ed. K Michaelian). Oxford, UK: Oxford University Press.
55. Tonegawa S, Liu X, Ramirez S, Redondo R. 2015 Memory engram cells have come of age. *Neuron* **87**, 918–931. (doi:10.1016/j.neuron.2015.08.002)
56. Guskjolen A, Cembrowski MS. 2023 Engram neurons: encoding, consolidation, retrieval, and forgetting of memory. *Mol. Psychiatry* **28**, 3207–3219. (doi:10.1038/s41380-023-02137-5)
57. Refaeli R, Kreisel T, Groysman M, Adamsky A, Goshen I. 2023 Engram stability and maturation during systems consolidation. *Curr. Biol*. **33**, 3942–3950.(doi:10.1016/j.cub.2023.07.042)
58. Robins S. 2023 The 21st century engram. *Wiley Interdiscip. Rev. Cogn. Sci*. **14**, e1653. (doi:10.1002/wcs.1653)
59. Xue G. 2022 From remembering to reconstruction: the transformative neural representation of episodic memory. *Prog. Neurobiol*. **219**, 102351. (doi:10.1016/j.pneurobio.2022.102351)
60. Heinen R, Bierbrauer A, Wolf OT, Axmacher N. 2023 Representational formats of human memory traces. *Brain Struct. Funct.* **229**, 513–529. (doi:10.1007/s00429-023-02636-9)
61. Favila SE, Lee H, Kuhl BA. 2020 Transforming the concept of memory reactivation. *Trends Neurosci*. **43**, 939–950. (doi:10.1016/j.tins.2020.09.006)
62. Ross TW, Easton A. 2022 The hippocampal horizon: constructing and segmenting experience for episodic memory. *Neurosci. Biobehav. Rev.* **132**, 181–196. (doi:10.1016/j.neubiorev.2021.11.038)
63. Zheng J *et al*. 2022 Neurons detect cognitive boundaries to structure episodic memories in humans. *Nat. Neurosci*. **25**, 358–368. (doi:10.1038/s41593-022-01020-w)
64. Leutgeb S, Leutgeb JK, Treves A, Moser MB, Moser EI. 2004 Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* **305**, 1295–1298. (doi:10.1126/science.1100265)
65. Schlichting ML, Zeithamova D, Preston AR. 2014 CA1 subfield contributions to memory integration and inference. *Hippocampus* **24**, 1248–1260. (doi:10.1002/hipo.22310)
66. Hainmueller T, Bartos M. 2020 Dentate gyrus circuits for encoding, retrieval and discrimination of episodic memories. *Nat. Rev. Neurosci*. **21**, 153–168. (doi:10.1038/s41583-019-0260-z)
67. Andermane N, Joensen BH, Horner AJ. 2021 Forgetting across a hierarchy of episodic representations. *Curr. Opin. Neurobiol*. **67**, 50–57. (doi:10.1016/j.conb.2020.08.004)
68. Ezzyat Y, Davachi L. 2011 What constitutes an episode in episodic memory? *Psychol. Sci*. **22**, 243–252. (doi:10.1177/0956797610393742)
69. Hanert A, Weber FD, Pedersen A, Born J, Bartsch T. 2017 Sleep in humans stabilizes pattern separation performance. *J. Neurosci*. **37**, 12238–12246. (doi:10.1523/JNEUROSCI.1189-17.2017)
70. Schapiro AC, McDevitt EA, Chen L, Norman KA, Mednick SC, Rogers TT. 2017 Sleep benefits memory for semantic category structure while preserving exemplar-specific information. *Sci. Rep.* **7**, 1–13. (doi:10.1038/s41598-017-12884-5)
71. Gilboa A, Moscovitch M. 2021 No consolidation without representation: correspondence between neural and psychological representations in recent and remote memory. *Neuron* **109**, 2239–2255. (doi:10.1016/j.neuron.2021.04.025)

72. O'Sullivan F, Ryan T. 2023 If engrams are the answer, what is the question? In *Engrams: a window into the memory trace*. Cham, Switzerland: Springer. (doi:10.31234/osf.io/f6amv)
73. Lengyel M, Dayan P. 2007 Hippocampal contributions to control: the third way. *Adv. Neural Inf. Process. Syst.* **20**, 889–896.
74. Blundell C, Uria B, Pritzel A, Li Y, Ruderman A, Leibo JZ, Hassabis D. 2016 Model-free episodic control (doi:arXiv:1606.04460)
75. Pritzel A, Uria B, Srinivasan S, Badia AP, Vinyals O, Hassabis D, Blundell C. 2017 Neural episodic control (eds D Precup, YW Teh). In *Proceedings of the 34th International Conference on Machine Learning*, vol. 70, pp. 2827–2836, Proceedings of Machine Learning Research. https://proceedings.mlr.press/ v70/pritzel17a.html.
76. Zeng X, Diekmann N, Wiskott L, Cheng S. 2023 Modeling the function of episodic memory in spatial learning. *Front. Psychol*. **14**, 1160648. (doi:10.3389/fpsyg.2023.1160648)
77. Kumaran D, McClelland JL. 2012 Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychol. Rev.* **119**, 573–616. (doi:10.1037/a0028681)
78. Zhou Z, Singh D, Tandoc MC, Schapiro AC. 2023 Building integrated representations through interleaved learning. *J. Exp. Psychol*. *Gen.* **152**, 2666–2684. (doi:10.1037/xge0001415)
79. Tulving E, Watkins MJ. 1975 Structure of memory traces. *Psychol. Rev*. **82**, 261–275. (doi:10.1037/h0076782)
80. Werning M. 2020 Predicting the past from minimal traces: episodic memory and its distinction from imagination and preservation. *Rev. Philos. Psychol*. **11**, 301–333. (doi:10.1007/s13164-020-00471-z)
81. Østby Y, Walhovd KB, Tamnes CK, Grydeland H, Westlye LT, Fjell AM. 2012 Mental time travel and default-mode network functional connectivity in the developing brain. *Proc. Natl Acad. Sci.* **109**, 16800–16804. (doi:10.1073/pnas.1210627109)
82. Nyhout A, Mahy CEV. 2023 Episodic thought in development: on the relation between memory and future thinking. *Dev. Rev.* **70**, 101103. (doi:10.1016/j.dr.2023.101103)
83. Sutton J, O'Brien G. 2022 Distributed traces and the causal theory of constructive memory. In *Current controversies in philosophy of memory* (eds A Sant'Anna, CJ McCarroll, K Michaelian), pp. 82–104. London, UK: Routledge.
84. Francken JC, Slors M, Craver CF. 2022 Cognitive ontology and the search for neural mechanisms: three foundational problems. *Synthese* **200**, 378. (doi:10.1007/s11229-022-03701-2)
85. Cowell RA, Barense MD, Sadil PS. 2019 A roadmap for understanding memory: decomposing cognitive processes into operations and representations. *eNeuro* **6**, ENEURO.0122-19.2019. (doi:10.1523/ENEURO.0122-19.2019)
86. Ferbinteanu J. 2019 Memory systems 2018—towards a new paradigm. *Neurobiol. Learn. Mem*. **157**, 61–78. (doi:10.1016/j.nlm.2018.11.005)