

Title: What is Episodic Memory if it is a Natural Kind?

Running Title: What is Episodic Memory?

Authors and author addresses

Prof. Dr. Sen Cheng
Mercator Research Group "Structure of
Memory"
Department of Psychology
Ruhr University Bochum
Universitätsstr. 150
44801 Bochum
Germany
Office Phone: +49 (234) 32 27136
Fax: +49 (234) 32 14463
Email: sen.cheng@rub.de

Prof. Dr. Markus Werning
Department of Philosophy II
Mercator Research Group "Structure of
Memory"
Ruhr University Bochum
Universitätsstr. 150
44801 Bochum
Germany
Office Phone: +49 (234) 32 24734
Fax: +49 (234) 32 14463
Email: markus.werning@rub.de

Corresponding author: Sen Cheng

Number of figures: 1

Number of words: 234 (abstract), 16010 (text), 404 (appendix), 20470 (text + appendix, references)

Acknowledgements: We thank Thomas Suddendorf for helpful discussions and Kevin Reuter for comments on the manuscript. This work was supported by a grant (SFB 874, project B2) from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and a grant from the Stiftung Mercator.

What is Episodic Memory if it is a Natural Kind?

Abstract:

Colloquially, episodic memory is described as "the memory of personally experienced events". Even though episodic memory has been studied in psychology and neuroscience for about six decades, there is still great uncertainty as to what episodic memory is. Here we ask how episodic memory should be characterized in order to be validated as a natural kind. We propose to conceive of episodic memory as a knowledge-like state that is identified with an experientially based mnemonic representation of an episode that allows for a mnemonic simulation thereof. We call our analysis the Sequence Analysis of Episodic Memory since episodes will be analyzed in terms of sequences of events. Our philosophical analysis of episodic memory is driven and supported by experimental results from psychology and neuroscience. We discuss selected experimental results that provide exemplary evidence for uniform causal mechanisms underlying the properties of episodic memory and argue that episodic memory is a natural kind. The argumentation proceeds along three cornerstones: First, psychological evidence suggests that a violation of any of the proposed conditions for episodic memory amounts to a deficiency of episodic memory and no form of memory or cognitive process but episodic memory fulfills them. Second, empirical results support a claim that the principal anatomical substrate of episodic memory is the hippocampus. Finally, we can pin down causal mechanisms onto neural activities in the hippocampus to explain the psychological states and processes constituting episodic memory.

Keywords: natural kind, neural sequences, replay, mental time travel, memory trace, events, reliable process, mnemonic simulation, autooensis

1. Introduction

Episodic memory is widely studied in neuroscience, psychology, and maybe less so in philosophy. It is a critical part of the human mind and has frequently been claimed to be a cornerstone of personal identity. Yet, there is no universal consensus on what constitutes episodic memory. In many textbooks, the notion of episodic memory is introduced in a classical taxonomical manner – *per genus proximum et differentiam specificam* (a classification scheme attributed to Aristotle): In a first step, a distinction between explicit and implicit memory (Graf and Schacter 1985), or similarly between declarative and non-declarative memory (Squire and Zola-Morgan 1988) is made. In a second step, two subordinate categories are introduced within the superordinate category of declarative memory, namely, semantic memory and episodic memory. This distinction was first introduced by Tulving (1972). Episodic memories were initially proposed to be those of personally experienced episodes, such as "I met my wife on my first day at work at Acme Co." By contrast, semantic memories were taken to consist of knowledge about the world such as "Abraham Lincoln was assassinated at Ford's Theatre in 1865". To elaborate on that content-based distinction, Tulving introduced the what-where-when (WWW) criterion to define the content of episodic memory (Tulving 1972). However, this criterion was found to be insufficient to distinguish semantic from episodic memories. As a result, Tulving later revised his definition of episodic memory based on autooetic consciousness, the conscious reliving of a past experience (Tulving 1985). Suddendorf and Corballis (1997) went even further and suggested that episodic memory is linked to mental time travel into the past and facilitates mental time travel into the future.

In this paper, we digress from this classical taxonomical pathway towards ontological classification. We will, instead, focus on the question of whether episodic memory is a natural kind and what implications this has for what episodic memory is best taken to be. It has recently become a central topic in the philosophy of psychology to ask whether certain notions used in psychology correspond to natural kinds and how this might assure that psychology has the inductive and explanatory potential we generally expect from sciences (Machery 2009). The question whether memory, in general, is a natural kind has been addressed before by Michaelian (2010) who argues for a negative answer. Furthermore, Bedford (1997) has argued that implicit memory is the result of a fallacy and should not be considered a category of memory. In this paper, we aim at expressing and consolidating our optimism that episodic memory, indeed, is likely to be a natural kind. The issue of whether semantic memory is a natural kind will not be addressed here.

For systematic reasons, we regard this project as an inherently interdisciplinary one, as a project that crucially depends on both philosophical and neuroscientific issues – one author of this paper is a philosopher, the other a theoretical neuroscientist. The enterprise is imperiled by two cliffs: A philosophical analysis of episodic memory and concepts related to it, even if such an analysis would be in good accordance with our conceptual intuitions, linguistic practices, or perhaps an introspective phenomenology of sorts, may well fail to classify a phenomenon that constitutes a natural kind. The question that becomes relevant when we are looking for a natural kind – for an exposition of our notion of a natural kind see section 4 – is the question of whether an underlying uniform causal mechanism exists. This question can only be settled with the help of the natural sciences. A concept shaped by philosophical analysis alone, however well this analysis is done, may still lack a corresponding natural kind and thus a referent that can be regarded as real from the point of view of scientific realism (Boyd 1991). This is the first cliff to be avoided. An inverse peril sticks out on the side of neuroscience and psychology. The description of particular brain mechanisms in neuropsychological memory research under the label of “episodic memory” may be regarded as a change of subject unless it can be demonstrated that the mechanisms provide a uniform basis for what we understand to be episodic memory. Objections of the sort “You have changed the subject” have a long tradition in the philosophy of mind when it comes to claims of identity or reducibility between mental and physical properties (Ryle 1949; Bieri 1995).

To show a way to navigate between the two cliffs is the main purpose of this paper. We are optimistic because we think that, on the one hand, we can provide a convincing philosophical analysis of episodic memory as experientially grounded representation with the mnemonic content of an episode and the potential for mnemonic simulation. On the other hand, in the light of recent neuroscientific findings, we propose that the principal anatomical substrate of episodic memory is the hippocampus. We suggest mechanisms of temporal sequence encoding, storage and retrieval in the hippocampus that are likely to constitute uniform causal mechanisms that underlie experientially grounded representations with episodes as mnemonic content and enable mnemonic simulations of those episodes. We thus would have made the case that episodic memory is a natural kind.

We begin with a critical review of prominent alternative accounts of episodic memory in section 2. Guided by a number of desiderata, we, in section 3, propose a new account of episodic memory, which we call the Sequence Analysis of Episodic Memory. Section 4 takes us from our central proposition that a notion of episodic memory should be such that it refers to a natural kind to an exposition of the cornerstones of our argumentation. These cornerstones set the agenda for the three subsequent sections: In section 5 we argue that the Sequence Analysis is both minimal and maximal with regard to its inductive and explanatory potential. Section 6 substantiates the claim that the principal anatomical substrate of episodic memory is the hippocampus. In section 7 we discuss evidence that neural processes in the

hippocampus provide uniform causal mechanisms for the processing stages proposed by the Sequence Analysis. We conclude that episodic memory – as characterized by the Sequence Analysis – can justly be expected to be a natural kind.

2. Different views on episodic memory

2.1. *The criterion based on content: what-where-when*

When Tulving (1972) first introduced the distinction between episodic and semantic memory, he suggested that episodic memories were unique in that they included information about the what-where-when of an event. Importantly, the three different types of information have to be represented jointly in a single memory, not separately in different memories (Clayton et al 2003). Since the what-where-when criterion refers to the content of a memory that can be tested in non-human animals, it has been frequently employed in animal cognition studies. Memory of joint what-where-when information of an event has been reported in food-caching scrub jays (Clayton and Dickinson 1998), pigeons (Zentall et al 2001), mice (Dere et al 2005) rats (Babb and Crystal 2005), chimpanzees and bonobos (Martin-Ordas et al 2010). While what-where-when information is undoubtedly frequently part of the content of episodic memory, the WWW criterion is insufficient to fully distinguish episodic from semantic memory since semantic memories can also include what-where-when information (Tulving 1985).

Memories that satisfy the WWW criterion in animal cognition studies are often referred to as episodic-like memories (Clayton and Dickinson 1998). In addition to the WWW criterion, some authors have suggested that for memories to be episodic-like they must be used flexibly, not only for rigid behaviors (Clayton et al 2003). Such flexible behavior contradicts the Bischof-Köhler hypothesis, so-named by Suddendorf and Corballis (Suddendorf and Corballis 1997), which states that nonhuman animals can react only to immediate needs and cannot take actions for future needs that conflict with their current needs (Köhler 1925; Bischof 1980; Bischof-Köhler 1985). To avoid this contradiction, recent studies in animal cognition in this area have sought to refute the Bischof-Köhler hypothesis (Naqshbandi and Roberts 2006; Raby et al 2007; Paxton and Hampton 2009). Most importantly, it has been shown that scrub jays, whose caching behavior meets the WWW criterion, could plan for a future need independently of their current needs (Raby et al 2007; but see Suddendorf and Corballis 2008). After having been trained that they receive food in the morning only in one room, but not in another, scrub jays were given free access to food to cache. Perhaps not surprisingly, the majority of the food was cached in the room without breakfast. However, in contradiction to the Bischof-Köhler hypothesis, the scrub jays were fed before the caching trial and therefore had no immediate need to cache food.

In our view, the what-where-when criterion is insufficient to characterize episodic memory. In some cases, the WWW criterion is too rigid. For instance, in episodic memories, one of the WWW components can be poorly encoded or missing (Friedman 1993; Bauer et al 2012). In other cases, the WWW criterion is too broad. For instance, semantic memory of an event that was not personally experienced may also contain all WWW components (Tulving 1985). To distinguish such semantic memories from episodic memories, Clayton et al. (2003) have suggested to add two more conditions, the structural and the flexibility condition. However, even with these additions Clayton et al. acknowledge that the combined criteria describe episodic-like memory, not proper episodic memory. In summary, even if the WWW criterion is convenient in nonhuman animal studies (Clayton and Dickinson 1998; Morris 2001) and supplemented by other conditions (Clayton et al 2003), it does not appear to appropriately capture the very idea of episodic memory in humans.

2.2. *Descriptive approaches*

An alternative approach to analyze episodic memory is to list the properties of episodic memory and show that the sum of these properties distinguishes it from other forms of memory, in particular semantic memory. This approach, too, has been pioneered by Tulving who proposed that semantic memory differs from episodic memory in 28 properties (Tulving 1983, Table 3.1, p. 35). Episodic memories are memories of events or episodes that are organized temporally. By contrast, semantic memories are memories of facts, ideas, or concepts that are organized conceptually. However, the distinction is not as clear as it may seem at first glance since information about an event or episode that occurred is information about a fact, too. Some properties remain vague, such as the temporal organization of the information in episodic memory, a notion which plays a central role in our analysis below. Other properties are properties of the neural system supporting memory, rather than of individual memories, such as appearing early or late in development. It is conceivable that the sum of all 28 properties would clearly distinguish between semantic and episodic memory. However, it remains unclear which properties are important characteristics of episodic memory and which ones are secondary consequences of those properties. Tulving later shifted his attention to focus on the difference in the subjective experience of memory retrieval, which we discuss below (Tulving 1985).

More recently, other authors have followed a similar approach. Conway (2009, Table 1, p. 2306) suggested nine different properties of episodic memories, and Henke (2010) suggested that memories can be distinguished along three dimensions: the number of learning trials required for memory formation, the cognitive complexity of the memory (single item vs. associative encoding), and the flexibility of the mental representation (Henke 2010). In this space of memories, episodic memories are rapidly encoded and flexibly represented associations, whereas semantic memories are slowly encoded and rigidly represented associations. However, more recent experiments have shown that semantic learning can occur rapidly after a single learning trial (Sharon et al 2011), calling Henke's account into question.

2.3. *The criterion based on subjective experience during retrieval: mental time travel*

Having realized the difficulty in distinguishing episodic memory from semantic memory on the basis of content alone, Tulving (1985) instead suggested a criterion based on the subjective experience during retrieval:

“Semantic memory is characterized by noetic consciousness. Noetic consciousness allows an organism to be aware of, and to cognitively operate on, objects and events, and relations among objects and events, in the absence of these objects and events. [...]

Of special interest in the present paper is auto-noetic consciousness, correlated with episodic memory. It is necessary for the remembering of personally experienced events. When a person remembers such an event, he is aware of the events as a veridical part of his own past existence.” (Tulving 1985)

Tulving further described the difference between auto-noetic and noetic consciousness as the difference between a person remembering a particular episode and knowing some information about the episode. The experience during retrieval of episodic memories was likened to mental time travel into the past (Tulving 1985; Tulving 1993), a reliving of the past. However, the purpose of traveling back in time is to inform future behavior. Therefore, the episodic memory system was suggested to sustain mental time travel into the future (Tulving 1985; Suddendorf and Corballis 1997). This suggestion is supported by experimental studies that show that amnesics have deficits in constructing imaginary scenes (Hassabis et al 2007)

and that the hippocampus is activated when healthy subjects imagine a future event (Weiler et al 2010).

Appealing to the Bischof-Köhler hypothesis, Suddendorf and Corballis (1997) made a strong argument that mental time travel is unique to humans. Since the publication of this article, a number of studies, some of which we mentioned in the previous subsection, have sought to prove the Bischof-Köhler hypothesis wrong. Suddendorf and Corballis, however, are not convinced by these studies because of methodological concerns (Suddendorf and Corballis 2008) and because the studied behaviors are limited to ecological niche behaviors, such as food caching in scrub jays (Suddendorf and Corballis 2007). Planning for the future in a narrow behavioral context is not equivalent to mental time travel into the future.

Although Suddendorf and Corballis originally proposed the mental travel idea to study foresight, it is arguably the currently most widely used account of episodic memory in humans. Despite its popularity, we believe that this approach is unsatisfactory at several levels. At a practical level, mental time travel is difficult, if not impossible, to study in nonhuman animals (Morris 2001; Clayton et al 2003) since it relies on the subjective experience during retrieval, which perhaps cannot be shared between different species. While it may turn out in the end that nonhuman animals do not possess episodic memory, that conclusion would be much stronger if episodic memory was not construed in such a way as to preclude, or severely bias against, this possibility. At a theoretical level, it seems unsatisfactory that the nature of a memory would depend predominantly on the subjective experience during recall since the memory persists even when not being recalled. One possible remedy might be to specify that a memory is episodic memory if an auto-noetic recall could be cued (Klein 2013). However, at a conceptual level, doubts are emerging about the association between different levels of consciousness and the retrieval of certain memories. For instance, it has been suggested that there is no clear dissociation between explicit and implicit memory systems (Berry et al 2008; Dew and Cabeza 2011) and that consciousness does not always accompany episodic memory retrieval (Hannula and Ranganath 2009; Henke 2010).

2.4. *Summary*

Each of the accounts discussed above focuses on certain aspects of episodic memory, while neglecting others. The what-where-when criterion focuses on the content of the memory, the descriptive approaches on a set of apparent properties, and the mental time travel hypothesis on the subjective experience during retrieval. Each approach has its strengths and drawbacks, accounting for some experimental results, but not for others. In the following, we propose the Sequence Analysis of Episodic Memory, which, while containing some elements of the previous approaches, rejects other elements from each of them. The Sequence Analysis is therefore not a synthesis of these previous approaches.

3. Episodic memory – a philosophical analysis

Our intention is to develop a philosophical analysis of episodic memory that is firmly grounded in empirical research from psychology and neuroscience. To set the stage for a rigorous analysis, we first sketch some desiderata to account for the most important empirical properties of episodic memory:

- (D1) The analysis needs to clarify what is potentially stored in episodic memory, that is, its content (discussed in section 3.1).
- (D2) Despite subjective experiences of recalling detailed episodic memories, numerous experimental studies have consistently found that episodic memory in humans often

preserves little more than the gist of the experienced episode. Therefore, our analysis has to integrate two competing requirements. On the one hand, the memory of an episode E must be allowed to differ in content (even significantly) from the experience of the grounding episode E^* . On the other hand, our analysis has to enforce a sufficiently stringent relationship between the experiential base E^* and the mnemonic content E to justify that the memory is based in the experience (discussed in 3.3).

- (D3) Overwhelming evidence indicates that subjects frequently retrieve inaccurate information when asked to recall episodic memories (see section 5.1). We regard these cases as improper episodic memory and aim for an analysis of episodic memory that presupposes its factivity¹ (discussed in section 3.5). Such a strong notion of episodic memory, we think, is theoretically more adequate and, at the same time, has the practical advantage that it provides an incentive for empirical investigations of when and how the episodic memory system fails to yield proper episodic memory.
- (D4) The philosophical analysis should be in accordance with our knowledge of neural mechanisms that underlie episodic memory (discussed in sections 6 and 7). Only then we can hope that the analysis captures a natural kind.

The key to fulfilling these desiderata, we propose, is to emphasize the sequential nature of episodic memory. In our view sequentiality is a distinguishing structural feature of both the content of episodic memory and its underlying neural realization. Based on numerous experimental studies on the neural realization of episodic memory, we recently suggested that neuronal sequences play a crucial role for episodic memory (Cheng 2013; Cheng and Werning 2013). However, we were by no means the first to do so, since others before us have noted this property (Kahana et al 2008; Conway 2009) and some computational models have accounted for episodic memories as neuronal sequences (Levy 1996; Lisman 1999; Hasselmo 2012).

3.1. Definition of an episode

We propose a recursive definition of an episode. We thereby attempt to give an account of what the potential contents of episodic memories are. We are, however, aware that specifying the potential contents of episodic memories does not suffice to characterize episodic memory and to demarcate it against semantic memory. Our definition of an episode presupposes that events – being spatially and temporally extended concrete particulars – are elements of our ontology and that an (at least) partial order of proper temporal succession $<$ can be defined on them: $e < e'$ means that the event e occurs before the event e' .

¹ In psychology, the adjective-noun combination “false memory” is often used to refer to a false mnemonic representation. The use of this phrase by psychologists is sometimes interpreted as if memory in psychology would not be regarded as factive. However, this conclusion would be justified only if “false” were an intersective adjective, for which the inference from “ x is AN” to “ x is N” is valid. The more plausible interpretation, we think, is that “false” in “false memory” is a privative adjective like “false” in “false money” or “fake” in “fake gun”. For privative adjectives, the inference is not valid: false money is not money, a fake gun is not a gun and, likewise, false memory is not a case of memory. Moreover, in psychological research and, even more so, in forensic situations the question of whether a memory report of the form “I remember that ...” truly is a case of memory or rather a case of confabulation or error often arises and is naturally regarded as a sensible question. This question would be pointless if memory were not generally regarded as factive.

Definition:

- (A1) If e_1 and e_2 are (primitive)² events and if $e_1 < e_2$, then the ordered pair $\langle e_1, e_2 \rangle$ is an episode.
- (A2) If $\langle \dots, e_i \rangle$ is an episode and e_{i+1} is an event and if $e_i < e_{i+1}$ then $\langle \langle \dots, e_i \rangle, e_{i+1} \rangle$ is an episode.
- (A3) Every episode is a (complex) event.

Following a standard convention, embedded brackets will be dropped: $\langle \langle \dots, e_i \rangle, e_{i+1} \rangle = \langle \dots, e_i, e_{i+1} \rangle$.

We regard our definition of an episode as minimal because further criteria to hold among the events of an episode could arguable be added: such as (i) temporal proximity, (ii) spatial proximity, (iii) causal structure (e.g., one event being the cause of another) or (iv) internal cohesion (e.g. shared participants). Sticking with the minimal definition, we do not want to preclude that two otherwise completely unrelated events – think of a child’s birth and a supernova that occurred millions of years ago, but is only perceived at about the same time – become part of the same episode in episodic memory³. Moreover, we do not even want to imply that every episode in the sense defined here could serve as content of episodic memory under realistic circumstances. Being episodically memorable may impose additional conditions on an episode that follow from the nature of episodic memory and its underlying mechanism.

By saying that an event is a particular (rather than a universal) we mean that it occurs only once: It neither repeats in time, nor does it occur as a whole at a different place at the same time. In the sentence “John reads a book every Sunday” the verb “read” does not denote an event. It rather denotes a class of events. In contrast, “read” in “John read his favorite book last Sunday” does denote an event. By assuming that events are concrete (rather than abstract), we mean that each event occupies a distinctive region in space-time, which no other therefrom independent event occupies. Spatial locations arranged in time may therefore serve as indices for events. For a more detailed analysis of the ontological status of events, see Pianesi and Varzi (2000).

Events may have a number of participants as in “Yesterday, John donated his favorite book to his parents” – i.e., John, the parents and the book. But need not as in “Yesterday it was raining”⁴. We are also inclined to accept rather sparse events as in “Last weekend, John was there”. Even though we often use linguistic expressions to refer to events, we do not want to tie our account of events to any particular representational format and in particular not to language. While many discussions of events focus on language (Davidson 1980; Parsons 1990; Trustwell 2011), we do not want to confine the theory of events to semantics.

For our proposed analysis, what counts is the relationship between episodes and events. Our definition of episodes implies that an episode is an ordered list of events. As such, an episode is distinct from a set of events, which is an unordered list of events.

² Our definition presupposes some version of event atomism, i.e., the assumption that there are primitive events that do not further divide into other events. We, however, leave open on which ontological level those events occur (on a microphysical or rather on a neuro-cognitive level). We regard this issue as inessential to the following argumentation of the paper.

³ There is experimental evidence that humans segment their experience into distinct episodes. For instance, Ezzyat and Davachi (2011) reported that cued retrieval is more successful if cue and target were perceived to have occurred within the same episode as opposed to distinct episodes.

⁴ For a neurophilosophical account of events and their participating objects, see Werning (2003).

Example: [John has dinner] is an episode because it is the sequence of events [John sits down at his dining table] < [John drinks red wine] < [John eats a tomato soup] < [John eats a steak] < [John drinks coffee]. By contrast, the set of these events would not constitute an episode, even though the unordered list of events would still make it true that John was having a meal. Since characterizing content of episodic memory is insufficient to fully constrain episodic memory, other conditions have to be satisfied.

3.2. Sequence Analysis of Episodic Memory

We continue our investigation into whether episodic memory is a natural kind with the following more explicit analysis of the concept of episodic memory. For this analysis, sequential representations in both experience and memory are crucial. We therefore call this analysis the Sequence Analysis of Episodic Memory. The question, discussed later on, will be whether there is a uniform neural mechanism that is a good candidate for the realization of episodic memory so characterized:

A subject S has episodic memory with content E at a time t_1 if and only if the following conditions are fulfilled:

- (S1) E is an episode with $E = \langle e_1, \dots, e_n \rangle$. E is called the *mnemonic content*.
- (S2) At some time t_1 , S compositionally represents E as an episode of temporally succeeding events e_1, \dots, e_n . S 's representation of E at t_1 is called the *mnemonic representation*.
- (S3) At a time $t_0 < t_1$, S has a reliable experience of the temporally succeeding events e_1^*, \dots, e_m^* , which make up an episode $E^* = \langle e_1^*, \dots, e_m^* \rangle$. E^* is called the *experiential base*.
- (S4) The episode E^* occurs at or before t_0 (*factivity*).
- (S5) The mnemonic content E is *ontologically grounded* in the experiential base E^* in the following sense of counterfactual dependence: Were E^* to occur at or before t_0 , E would also occur at that time.
- (S6) S 's representation with content E at t_1 is *causally grounded* in S 's experience of E^* through a reliable memory trace.
- (S7) On the basis of its mnemonic representation with content E , S is capable of generating a temporally explicit simulation with content E at some time $t_2 \geq t_1$. The generated simulation is called a *mnemonic simulation*.

These conditions can be related to the four major stages of memory processing: perception, encoding, storage and retrieval. (S3) and (S4) propose conditions on perception, (S5) and (S6) on encoding, (S1) and (S2) on storage, and (S7) on retrieval. Mapping conditions to processing stages may help translate between the more philosophical terminology of the Sequence Analysis and the terminology used in psychology and neuroscience.

3.3. The relation between experiential base and mnemonic content

Episodic memory is grounded in experience. A major challenge for any analysis of episodic memory is to characterize the relation between the content (E) of episodic memory and the content (E^*) of the experience it is based on. Had I not experienced the piano concert performed by the famous pianist in our concert hall last Saturday, I would not remember the melody of the sonata today. However, the content (E^*) of my experience then was very *different*, in a sense, from the content (E) of my memory now. The experience involved more

than just the auditory sense. I experienced the concert perhaps with all my senses, but at least also with the visual one. My memory today is merely auditory. Furthermore, even the quality of the auditory content is very different: The auditory content of the experience was much richer, more nuanced, detailed and vivid than the content of my memory is now. Still, my memory now (i.e., the mnemonic representation of E at t_1) and my experience then (i.e., the experience of E^* at t_0) are not representations of two *distinct* concerts. What I experienced then *makes true* what I remember now. Otherwise my mnemonic representation would be deficient. I would have a mnemonic representation of something that I did not experience.

One proposal to capture the dependence of mnemonic content on the experiential base would be to appeal to a logical entailment relation between them (for a discussion see Bernecker 2010). However, this would make appeal to the rather presumptuous premise that both the representation of the experiential base and the representation of the mnemonic content are conceptual. For, logical entailment relations can only hold between conceptual (or linguistic) representations. At least for the experiential representations this is highly controversial (Raffmann 1995; Bermudez 1995; Kelly 2001; Toribio 2007). This is why we suggest a counterfactual dependence relation as expressed by condition (S5). Given Lewis's (Lewis 1973) standard account of counterfactuals in terms of possible worlds, condition (S5) amounts to the following: Every close possible world in which the experiential base E^* occurs at or before t_0 , is a world in which the mnemonic content E also occurs at that time. In other words, the experienced occurrence of the episode E^* secures the occurrence of the remembered episode E . E is not just contingently linked to E^* , but grounded therein. We could also say: What one has experienced is a *truth-maker* of what one remembers. Our formulation warrants a sufficiently strong dependence relation without requiring the identity of experiential base and mnemonic content.

Mnemonic content can be ontologically grounded in an experiential base in a number of ways: (i) Identity: A probably merely theoretical option of ontological groundedness is the identity between mnemonic content and the experiential base. In that case the remembered episode would be identical to the experienced episode with regard to the events in the episode, the participants of these events, and the features of the events and participants. This is perhaps linked to the idea that some people have in mind when they talk about "photographical memory". (ii) Constituency: A more realistic option is that the mnemonic content is ontological grounded in an experiential base because the former is a part of the latter. In memory, some events of the experienced episode, some participants of the events, or some features of the events and participants might have been dropped. A person might have experienced how a dark-haired girl lost her vanilla ice cream and dropped it on her left black shoe after a tall boy with a yellow shirt jostled her. The person might just remember, though, that a girl dropped her ice cream after being jostled. (iii) Abstraction: A mnemonic content may also be ontological grounded in an experiential base when the former is an abstraction of the latter. This is the case if a very specific sequence of events, a single event or a participant or feature in the experienced episode is remembered as belonging to a more coarse-grained category; say, a sequence of kicks and punches on the body and the face as a physical attack, an assembly of seven sheep as a flock, or a particular shade of red as a warm color. Often psychologists speak of the gist of an episode in those cases.

3.4. *Mnemonic representation and mnemonic simulation*

In the Sequence Analysis we distinguish between an actual and temporally enduring mnemonic representation (S2) and a possible and only instantaneous mnemonic simulation (S7). Both are representations of the same episode E , which, however, have rather different representational formats. For an explicit characterization of the two formats see the appendix.

According to condition (S2), the mnemonic representation is a compositional representation in which the temporal succession of the events in the episode is encoded by some structure among the representational constituents. The encoding structure is typically not temporal itself and thus allows for an enduring representation of the episode. Compositionality – the principle that the content of a complex representation is a structure-dependent function of the contents of its parts – is a widely acknowledged, though not uncontroversial, criterion for the adequacy of representational structures in general, be they linguistic, conceptual or neural (Hodges 2001; Werning 2005; Pagin and Westerståhl 2010; Werning et al 2012).

Condition (S7) appeals to a potentiality insofar as it requires that the subject is capable of generating a temporally explicit simulation with content E (for an account of episodic memory as mental simulation see Shanton and Goldman 2010). The most salient difference between the enduring mnemonic representation of a temporal succession of events and its temporally explicit mnemonic simulation is that in the latter the temporal succession of events in the domain of representational contents is represented itself by a temporal succession of events in the domain of the representational vehicles – in our case neural processes (for simulation or emulation accounts of mental representation see Grush 2004; Werning 2012; Mroczko-Wąsowicz and Werning 2012). On the basis of empirical observations, it has been argued that the generative or constructive nature of the episodic memory system might be explained by postulating that information is added during retrieval (Bernecker 2008; Michaelian 2011a; Schacter 2012). It may, indeed, often be the case that in retrieval the subject generates a simulation of an episode \hat{E} , in which information – events, participants or features – is added to E such that \hat{E} contains, but is not identical to, E. Let us assume that the remaining conditions of the Sequence Analysis are fulfilled for the episode E, but not for the enriched episode \hat{E} , because the subject does not have a mnemonic representation of \hat{E} or did not even have an experience of \hat{E} . In such a case, the Sequence Analysis would entail that the subject does have episodic memory of E, but not of \hat{E} . This entailment holds whether or not \hat{E} actually occurred.

The requirement of a mnemonic simulation of E is in some respects akin to Tulving's and Suddendorf and Corballis' idea that episodic memory should allow the subject to "consciously relive" an episode. There is, however, an important distinction. Tulving (1985) regarded it on a priori grounds as essential to episodic memory that the "reliving of an episode" – the explicit simulation of the episode – is conscious. By contrast, the Sequence Analysis does not, a priori, presuppose that some form of consciousness is essential for episodic memory. However, we do not want to preclude that future research might result in an identification of mnemonic simulation and some form of consciousness or the establishment of a close link between them. Of course, much depends on a better understanding of the neural correlates of the various forms of consciousness.

3.5. *Factivity and the knowledge-likeness of memory*

The Sequence Analysis proposes an epistemically strong notion of episodic memory. First, our notion of episodic memory is factive: If S has episodic memory of the episode E, E in fact occurred. This follows from the conjunction of conditions (S4) and (S5). E* occurred at or before t_0 – due to (S4) – and if E* were to occur at or before t_0 , E would occur at that time – due to (S5). Second, episodic memory amounts to a knowledge-like state provided that knowledge is analyzed in reliabilist terms – i.e. as a reliably produced true belief (Goldman 1986). An information process is standardly regarded as reliable if it is conducive to truth with a probability greater than some threshold value. With regard to episodic memory, the reliability of the production process depends on two stages: At the stage of perception, condition (S3) warrants that the causal process leading to the grounding experience is reliable.

Regarding the stage of encoding, condition (S6) makes sure that the memory trace, too, is reliable.⁵

We are aware that in ordinary language the verb “remember” is sometimes also used in a non-factive way. In some contexts, speakers would be fine with a sentence like “John remembered that a green Renault jumped the red light, but in fact it was a grey Peugeot”. This occasional non-factive use is a linguistic feature that “remember” shares with many other mental attitude verbs like “to see”, “to hear”, and “to feel”. These are typically analyzed in a factive way such that “John saw that a green Renault jumped the red light” is taken to imply “A green Renault jumped the red light”. However, many speakers, in certain contexts, would also be fine with a non-factive use of the verb “to see” as in “When John had consumed LSD, he saw a pink elephant” or “John saw that a green Renault jumped the red light, but in fact it was a grey Peugeot”. The occasional, context-driven non-factive use does not undermine a general, context-neutral factive analysis. One only has to assume that particular contexts may coerce the interpretation of a word and may lead to a weakening of the word’s lexical component that is responsible for its factivity. Coercion is a wide-spread phenomenon in language.⁶

Note that a person’s having a false mnemonic representation and thus having improper episodic memory does not entail that the person’s episodic memory system is defective. The situation is somewhat analogous to certain cases of illusionary perception. Take for example the apparent motion illusion: On a screen, a red and somewhat distant green dot are shown alternatingly. Given a certain distance of the two dots and a certain frequency of their alternation, neurotypical subjects report seeing a moving dot that changes its color in the middle of the distance. It would, however, be false to say that any of these subjects actually sees a moving dot changing its color because none of the dots is actually moving. The subjects do not have a perception of a moving, color-changing dot, but a perceptual illusion thereof. This is so even though the subjects’ perceptual systems are functioning perfectly well. Ophthalmologists even sometimes use visual illusions to test whether a subject’s visual system is working properly.

The situation is analogous to that of episodic memory. Having a false mnemonic representation, e.g., due to an episodic memory illusion, is fully consistent with a subject’s episodic memory system working properly. Below we will discuss cases in which episodic memory illusions were explicitly induced in neurotypical subjects (section 5.1). But just as perceptual illusions in subjects with a properly functioning perceptual system are cases of improper perception, episodic memory illusions in subjects with a properly functioning episodic memory system are cases of improper episodic memory.

4. Is episodic memory a natural kind?

When we ask whether episodic memory is a natural kind, we presuppose a notion of natural kind that can be traced back to Richard Boyd (Boyd 1991; 1999). It is commonly labelled “the homeostatic property cluster view” (HPC view) of natural kinds. The core idea is that, in science, entities should be clustered together in a way that (i) optimizes the inductive and explanatory potential of theories that make reference to those clusters and (ii) that this

⁵ For a discussion on reliable memory traces, see (Martin and Deutscher 1966).

⁶ It is astounding that in some linguistic cases even the use of the noun “knowledge” does apparently not imply factivity. In the quite frequent phrase „his knowledge was outdated“ the predicate “outdated” coerces an interpretation of “knowledge” as denoting a set of propositions that are not true given the current state of evidence.

inductive and explanatory potential should rest on uniform causal mechanisms underlying each cluster. In the spirit of the HPC view and closely following Machery (2009), we will use the notion of natural kind as defined in the following way:

A class *C* of entities is a natural kind if and only if there is a large set of properties that subserve relevant inductive and explanatory purposes such that *C* is the maximal class whose members are likely to share these properties because of some uniform causal mechanism.

Along with the HPC view, our definition reveals its strengths when contrasted with two extremes: On the one hand, it opposes essentialism, i.e. the view that each kind is identified by a necessary property, its essence (“Gold is whatever has atomic number 79”). If one were to presuppose essentialism, there would probably be no kinds left anywhere other than in elementary physics. This would belie the inductive and explanatory potential of sciences like biology, geology and psychology where essential properties are notoriously hard to identify. On the other hand, the HPC view shies nominalism according to which there are only nominal kinds (“the set of solid objects in Paris with a mass below 13 kg”). Our notion of natural kind seems to capture just the right idea of clustering to account for the inductive and explanatory power we observe the various sciences to have. It makes sure that for each natural kind, there is neither a subset nor a superset with just the same inductive and explanatory potential. It finally assumes that for each natural kind there is a uniform causal mechanism that explains why the members of the kind are likely to share the set of properties in question. Uniformity requires that the underlying causal mechanisms in all instances are of the same type, but permits that the mechanisms in question are themselves complex and decompose into various partial mechanisms. For different kinds, the underlying causal mechanism may also be very different: For instance, we have the causal mechanism of genetic flow in the case of biological species, the electromagnetic forces in the case of chemical elements, and the causal mechanism of language acquisition that explains why the members of particular language communities are likely to share a grammar.

In the following, we will argue that episodic memory as analyzed above indeed is a natural kind. Our argumentation will proceed along three cornerstones.

- (C1) The Sequence Analysis is both minimal and maximal with regard to its inductive and explanatory potential.
 - (C1.1) It is minimal because any violation of one of the conditions will lead to a deficiency in episodic memory (section 5.1).
 - (C1.2) It is maximal because other forms of memory do not satisfy the conditions nor do other cognitive processes (section 5.2).
- (C2) The principal anatomical substrate of episodic memory is the hippocampus.
 - (C2.1) The principal function of the hippocampus is episodic memory. That is, all processes hosted by the hippocampus contribute to episodic memory (section 6.1).
 - (C2.2) Episodic memory is principally hosted by the hippocampus. That is: Even though episodic memory involves interactions with other cognitive processes, which are supported by a variety of brain regions, processes specific to episodic memory are hosted by the hippocampus (section 6.3).
- (C3) Neural processes in the hippocampus provide uniform causal mechanisms for the processing stages proposed by the Sequence Analysis.
 - (C3.1) The hippocampus provides a uniform causal mechanism that aligns the sequential representation of mnemonic content with the sequential representation of the experiential base (section 7.1 on phase precession and theta sequences)

- (C3.2) The hippocampus provides a uniform causal mechanism for the compositional mnemonic representation of episodes and their mnemonic simulation in retrieval processes (7.2 on replay).
- (C3.3) Interventions in the memory trace warrant that mnemonic representations are causally grounded in experiences (section 7.3 on disruption of systems consolidation).

We are aware that these cornerstones define an ambitious research agenda that cannot be treated comprehensively in this paper. We will nevertheless discuss selected experimental results that provide exemplary evidence for each of the cornerstones.

5. The properties of episodic memory according to the Sequence Analysis

Here we outline why we believe that the conditions for episodic memory in the Sequence Analysis are unique to episodic memory. Due to the large number of cognitive processes that can be distinguished, we are not able to discuss the entire range of processes. Instead, we will focus on a few prominent candidates and, in particular, we give examples for cases, in which mnemonic representations satisfy nearly all conditions, but fail to be proper episodic memories.

5.1. *Why the Sequence Analysis is minimal*

As a logical consequence of the Sequence Analysis, there are a number of ways a mnemonic representation of an episode can fail to be an instance of episodic memory. We discuss the empirical evidence for each logical possibility to show that all the conditions in the Sequence Analysis are necessary, i.e., the Sequence Analysis is minimal (C1.1). In cases, where a number, but not all, conditions in the Sequence Analysis are satisfied, we will refer to the mnemonic representation as improper episodic memories⁷.

(a) The mnemonic representation of an episode E may be false.

This would be the case if the episode E did not occur. A mnemonic representation may be false even if both the experience and the memory trace are reliable. Being a probabilistic notion, the reliability of the production process does not entail the truth of the resulting representation (also see section 3.5 above).

(b) The mnemonic representation may be a case of improper episodic memory because there is no grounding experience at all regardless of whether the mnemonic representation is true or false.

Loftus et al. (1995) have demonstrated that it is possible to induce a mnemonic representation of an episode that is not grounded in any experience. In the lost-in-the-mall experiment, adult subjects listened to four descriptions of a childhood event supposedly provided by relatives. One of these stories, about being lost in a shopping mall at age 5, was false. Nevertheless, the

⁷ In psychology, the terms “false episodic memory” or “false memory” are popular (Marsh et al 2008). We use the broader term “improper episodic memory” to indicate that a mnemonic representation fails to be proper episodic memory if one or more of the conditions (S1)-(S7) is violated, even in cases where the content of the mnemonic representation is veridical. In other words, false episodic memories are improper episodic memories, but the reverse is not true.

story did include factual information from the subjects' childhood such as a description of a mall that the family usually shopped at. When asked later, six out of 24 subjects said that they "remembered" the false episode. Subjects were even able to provide details about the false episode. Interestingly, even after being told that one narrative was false, five out of the six subjects did not identify the lost-at-the-mall episode as the false narrative.

(c) Regardless of whether the mnemonic representation is true or false, it may be improper episodic memory because the grounding experience is not reliable.

In this context, it is important to notice that we do not want to restrict the term experience to experience based on external sensory inputs, but are open to include other forms of experiences, e.g., proprioceptive experiences (e.g., pain) and introspective experiences (Werning 2010). Proprioceptive and introspective events may as well be elements of episodes and hence may sometimes mix with perceptually experienced events. Now a mnemonic representation of an episode, say an encounter with a black panther, may fail to be a case of episodic memory because the representation is based on a misperception, illusion or even a hallucination. A mnemonic representation of that episode will also fail to be a case of episodic memory if it is based on the imagination or confabulation of that encounter. In all these cases, condition (S3) is violated. However, a subject may also form a mnemonic representation of an introspective episode: "She remembers that she *imagined* an encounter with a black panther"; "She remembers that she *hallucinated* an encounter with a black panther". Here the represented episode is the imagination of the encounter with a black panther, or, respectively, its hallucination. Each of the two utterances may truthfully refer to an instance of episodic memory if the person has had a reliable introspective experience of her imagination, or, respectively, her hallucination, and the remaining conditions of the Sequence Analysis are satisfied.

What if the mnemonic representation of an episode is grounded in the testimony of another person who may be reliable as a witness and may have experienced the episode, say a car accident, himself? What if I have read about the car accident in a court file? What if I have seen an offline video recording of the accident? It is true that the testimony of an episode by a personal witness, a text or a video may establish a reliable informational link between a subject and the episode. Still, according to the Sequence Analysis, in none of the three cases the subject has episodic memory of the car accident. The reason is that the subject has not experienced the car accident (Cohen and Meskin 2004). It is noteworthy that the ordinary language use of the verb "remember" does not make the appropriate distinctions here: Sometime after I listened to a witness, read the court file or saw the video, it might be fair to say that I remember that the blue Mercedes violated the green Beetle's right-of-way. This mnemonic representation, however, does not qualify as episodic memory. Again, my remembering that I *heard/read/saw-on-video* that the blue Mercedes violated the green Beetle's right-of-way does qualify as episodic memory.

(d) A mnemonic representation of an episode may fail to be a case of episodic memory – even though the subject has had a reliable experience – because the causal link between the subject's experience and the mnemonic representation is not established through a reliable memory trace.

Many psychological studies into eyewitness testimony have studied the reliability of the memory trace and found that subsequent experiences can alter the content of a stored memory (retroactive inference). Examples of retroactive inference are imagination inflation and the misinformation effect (Marsh et al 2008). The latter effect was first demonstrated by Loftus and Palmer (1974). They showed subjects a film of a traffic accident. After answering questions about the details of the accident, the participants were split into different groups. One group of subjects was asked "About how fast were the cars going when they *hit*

[emphasis added] each other?“, while another group was asked "About how fast were the cars going when they *smashed* [emphasis added] each other?" The only difference between the two groups was that the word "hit" was substituted by “smashed” in the question. After a week, subjects were asked whether they had seen any broken glass, even though there was none in the film. The number of subjects answering in the affirmative was significantly larger in the second than in the first group. This and many other experiments (Marsh et al 2008) show that subsequent experiences can intervene in the memory trace linking the grounding experience to the mnemonic representation. When this occurs the memory trace might become unreliable.

(e) A mnemonic representation can be defective because the experiential base does not secure the mnemonic content, e.g., if events are ordered fallaciously or added up.

The case, when events are ordered fallaciously or added up to an episode that has not been experienced, is referred to as misattribution in psychology (Schacter and Dodson 2001). Such a case occurred, for example, after the bombing of a federal building in Oklahoma City in 1995 (Schacter 2002). A mechanic in a rental shop reported that he had seen the prime suspect, Timothy McVeigh, together with an accomplice, referred to as "John Doe No. 2". After an extensive search for this second suspect, the police determined, that John Doe No. 2 had visited the shop on his own one day after McVeigh did. In controlled experiments, such intrusions of one memory into another can be induced reliably (Lindsay et al 2004).

(f) A mnemonic representation of an episode may fail to be a case of episodic memory because the subject is unable to generate a mnemonic simulation with content E.

This case would occur if a person had a reliable experience and a reliable memory trace, but was unable to retrieve the memory for some reason. It is difficult to study this case empirically because memory retrieval, the very instrument that usually demonstrates that a subject has a mnemonic representation, is defective. How else can we ascertain the existence of the mnemonic representation? In the example of repressed memories, memories of traumatic events are thought to be repressed by the subject due to the pain they cause, but can be recovered through external interventions such as psychotherapy or hypnosis (Patihis et al 2014). However, the concept of memory repression and recovery is highly controversial. Many authors believe that recovered memories are false memories, often induced by therapists inadvertently during the recovery process (Loftus 1993). In controlled experiments, it can be shown that subjects are able to recall items that they could not recall during previous testing. This phenomenon is called reminiscence (Payne 1987), but it has been studied mainly with words and images and the relationship to episodic memory remains unclear. While the empirical basis for this case of improper episodic memory is unclear, the possibility of its existence is widely acknowledged and the subject of many psychological studies.

(g) Finally, the content of the mnemonic simulation could fail to be of the same content as that of the mnemonic representation, while all other conditions in the Sequence Analysis are satisfied.

This case may occur, if additional information available during the retrieval process changes the information retrieved from memory. For instance, suggestive questions can bias the report of a subject’s memory (Scoboria et al 2002). In the study of Loftus and Palmer (1974), discussed above, the subject in the smash group reported a higher speed (40.8 mph) than the subjects in the hit group (34.2 mph). This case is related to, but distinct from, the misinformation effect discussed in (d) above. In this case, the suggestive question biases the subject’s immediate response, whereas the misinformation effect refers to the fact that the suggestive question interferes with the memory trace such that subsequent memory retrieval becomes faulty.

In the discussion above, we reviewed a number of cases in which post-episode experience interfered with, i.e., had a distorting influence on, previously stored memories. However, we note that post-episode experience could be consistent with previously stored episodic memories and even improve on it by, e.g., adding detail. Michaelian (2011b) suggested recently that the post-episode addition of accurate information occurs far more often than the addition of inaccurate information. Whether an episodic memory that was modified post-episodically with accurate information still constitutes a case of episodic memory in accordance with condition (S6) depends on whether we can still maintain, with some right, that the mnemonic representation is causally grounded in the experiential base through a reliable memory trace.

In summary, there is rich empirical evidence for a correspondence between realistic example cases and logical possibilities for improper episodic memories. Any of these logical possibilities are opened up by one or more conditions in the Sequence Analysis not being fulfilled. We therefore conclude that all conditions in the Sequence Analysis are required, i.e., that the Sequence Analysis is minimal (C1.1).

5.2. Episodic memory in the Sequence Analysis is clearly distinct from other types of memory

It does not take much imagination to see that most types of memory do not satisfy the conditions for episodic memory in the Sequence Analysis. For instance, memory acquired during perceptual learning, classical conditioning, and conditioned taste aversion is not episodic memory because the content is not representable as a sequence of events (S1). Learned motor skills, e.g., running, are representable as a sequence of small movements and shorter sequences can be combined into longer ones. Even though some memories of motor sequences may be cases of episodic memories, many motor sequences of any complexity are usually not learned in a single trial, or through an experience thereof, and therefore do not constitute episodic memory. Our assumption that events are particulars, together with (S5) and (S6), imply that episodic memories have to be formed after a single experience, which is also called one-trial (or one-shot) learning. The other examples of memory mentioned above in this paragraph, except for the case of conditioned taste aversion, are also not usually acquired after single learning trials.

In psychology and neuroscience there is a vigorous debate about whether a dividing line exists between semantic and episodic memory (McKoon and Ratcliff 1986; Toth and Hunt 1999; Klein 2013), and if yes, where to draw it⁸. We believe that this uncertainty has led to two opposing views in the psychological literature on what happens to an episodic memory after it has been encoded. McClelland et al. (1995) suggested that an episodic memory remains an episodic memory however long ago it was encoded. By contrast, Nadel and Moscovitch (1998) propose that through repeated retrieval of episodic memory, memory traces of a different type, semantic memory, is established by extracting regularities from the content of the episodic memory. These two different views lead to different explanations of systems consolidation, which we discuss in the next section.

⁸ A major methodological hurdle in studying the potential dissociation between semantic and episodic memory is that the retrieval might always involve both types of memory (e.g., McCabe et al 2011). However, irrespective of whether an experimental paradigm can be developed to dissociate the retrieval process behaviorally, there could be a conceptual and neural difference between the two types of memory.

The Sequence Analysis provides several ways to distinguish semantic memory from episodic memory based on different criteria. One possibility is that the content of the semantic memory does not constitute an episode. Take for example the memory that Paris is the capital of France. However, other cases of semantic memory are more difficult to distinguish from episodic memory, for instance, when the content of episodic and semantic memories are similar. The statement "I remember that I received an airplane for Christmas when I was eight-years-old" may fail to refer to a case of episodic memory because (i) the person has just an associative mnemonic representation of the various objects referred to in the sentence; (ii) the memory is not causally grounded in the experience, but rather on a second-hand report; or (iii) in retrieving the information the person does not generate a simulation of the remembered episode. We note that it is irrelevant for our argument, whether the person is aware of the mnemonic simulation.

There might be memories of sequences that are based on multiple experiences of the sequence and that allow for some kind of mental simulation of the sequence, for instance, the memory of the sequence of notes acquired by hearing a certain musical piece multiple times. Some people might call these memories cases of semantic memory. However, it remains an open question whether such memories differ from paradigmatic cases of episodic memory categorically or by degree. This raises the general question whether the difference between semantic memory and episodic memory is categorical or gradual. To address this issue, future conceptual and empirical work is needed.

In summary, experimental results suggest that no other memory or cognitive process satisfies all the conditions for episodic memory in the Sequence Analysis and thus that the Sequence Analysis is maximal (C1.2).

6. The anatomical basis of episodic memory

While in principle a uniform causal mechanism for episodic memory could be distributed widely in the brain, the case for a uniform causal mechanism might be simpler to make if the mechanism were localized in a particular region. Here we suggest that the uniform causal mechanism for episodic memory is, in fact, located in the hippocampus because the hippocampus appears to be the principal anatomical substrate of episodic memory (C2). We argue in the following that damaging or removing the hippocampus severely impairs episodic memory without much impact on other cognitive functions (C2.1, see sections 6.1-6.3) and that no other brain region appears to play a similarly selective role in episodic memory (C2.2, see section 6.4). This is not to say that episodic memories are exclusively encoded, stored and retrieved by the hippocampus. On the contrary, we believe that the hippocampus is only part of a network that performs these functions and this network includes among other structures the neocortex (for a more detailed discussion, see 6.4). However, the hippocampus appears to play a principle role in this network with regard to episodic memory.

6.1. The role of the hippocampus in episodic memory

The first and most important hint that the hippocampus is involved in episodic memory was the case of patient HM. After both his hippocampi were removed in a surgery to control his epileptic seizures, he could no longer form new episodic memories (Scoville and Milner 1957). This condition is called anterograde amnesia. Intriguingly, HM did not suffer apparent impairments of most other cognitive functions such as language, perception and working memory (Scoville and Milner 1957). Over the years, these basic and many other observations have been confirmed in a number of hippocampal patients (Squire and Zola-Morgan 1988), although doubts have been voiced, too, which we review in the next section. Amnesics also

lose memories of past episodes, i.e., from the period before the hippocampal damage (retrograde amnesia). Memories from the remote past appear to be less affected than recently formed memories. This gradient of retrograde amnesia had been observed earlier after head trauma that did not involve permanent brain damage (Ribot 1881; Müller and Pilzecker 1900). The process by which episodic memories become less prone to disruption is known as systems consolidation. Hippocampal damage leads to graded retrograde amnesia, suggesting that the hippocampus is not only required during the encoding of episodic memories, but also during systems consolidation.

Why systems consolidation is necessary has been the subject of much debate. McClelland et al. (McClelland et al 1995) suggested that due to catastrophic interference a neural network cannot both store memories rapidly and stably over time. As a consequence, they proposed that the brain uses two complementary learning systems (CLS) to store episodic memories. They are first quickly stored in the hippocampus, where plasticity is rapid, and then gradually transferred to the neocortex, where memories are encoded more slowly, but also more stably. In the CLS model, episodic memories are stored in both hippocampus and neocortex but at different rates. This prediction is consistent with imaging studies that find activations above baseline in the hippocampus for retrieval of recent memories, but activity in neocortical regions for remote memories (Bontempi et al 1999). However, it is hard to conceive why the transfer process would last 15 years or more, as suggested by the temporal gradient of retrograde amnesia observed in amnesics whose brain damage was limited to the hippocampus (Squire and Alvarez 1995).

By contrast, the multiple memory trace (MMT) theory proposes that episodic memories are stored only in the hippocampus (Nadel and Moscovitch 1997). To account for graded retrograde amnesia, Nadel and Moscovitch postulate that during each retrieval a new copy, or trace, of the memory is created. The older the episodic memory, the more frequently it generally has been retrieved and the more traces of it exist. These multiple traces would then make the older memories less prone to (partial) lesions of the hippocampus. According to MMT theory, autobiographical memory that persists after complete hippocampal lesions is not episodic memories, but event information stored in semantic memory, which is supported by neocortex (Nadel and Moscovitch 1998; Cheng 2013). MMT theory therefore predicts that recall of episodic memory requires the hippocampus, no matter how remote the memory is. This prediction is supported by findings in amnesics (Steinvorth et al 2005) and fMRI studies (Nadel et al 2000; Ryan et al 2001; Harand et al 2012). We therefore conclude that storage and retrieval of episodic memories always require the hippocampus.

6.2. Other functional roles attributed to the hippocampus

While it was initially accepted that patient H.M. and other patients with hippocampal lesions have deficits only in episodic memory, later studies have found potential deficits in other cognitive functions. Some of these findings were corroborated by observations of hippocampal activation in healthy subjects performing those same tasks. We briefly summarize some findings of this kind in this section before proposing in the following section that these findings can generally be accounted for by an involvement of episodic memory in the cognitive tasks that were employed.

Perception of objects. The most common task used to study this link is a visual discrimination task. Patients with extensive lesions to the medial temporal lobe have deficits in discriminating between very similar visual stimuli (Lee et al 2005; Lee and Rudebeck 2010; Barense et al 2012).

Perception of space. Discrimination of images of spatial scenes is impaired in patients with focal lesions of the hippocampus (Lee et al 2005). There has been a proliferation of studies that link the hippocampus to visual perception (see Baxter 2009 for a review). During this task, fMRI studies find that, in control subjects, the hippocampus is activated during a visual discrimination task (Aly et al 2013).

Language. While language ability initially seemed unaffected in patient HM, careful investigations found that with advanced age HM had forgotten low-frequency words more often than age-matched controls (James and MacKay 2001). HM had more difficulty to remember both the meaning of low-frequency words and whether the presented text was in fact an English word or not.

Short-term memory. Amnesics have deficits in short-term-memory tasks (Aggleton et al 1992; Owen et al 1995; Holdstock et al 1995; Holdstock et al 2000), so do monkeys with lesions of the medial temporal lobe (Zola-Morgan et al 1989). BOLD activity in the medial temporal lobe is associated with active maintenance of novel information (Ranganath and D'Esposito 2001) and the degree of activation during the delay predicts later long-term memory performance (Ranganath et al 2005).

Temporal associations. The hippocampus is required for learning tasks that require associations across temporal gaps and processing of temporal sequences. For instance, in learning sequences of odors in the same location (Fortin et al 2002), disambiguation of overlapping sequences (Agster et al 2002) and for trace conditioning (Weiss et al 1999).

Semantic memory. Acquisition of new semantic memory was reported to be very slow and laborious in amnesics (Levy et al 2004). Klein (2013) recently proposed that semantic and episodic memory do not differ in their memory trace during storage, i.e., the nature of the stored memory trace is uncategorized until it is retrieved. Only during retrieval is the memory differentiated into semantic and episodic memory, as suggested by Tulving (1985), based on the type of consciousness that accompanies retrieval, noetic and auto-noetic, respectively.

Scene construction. Amnesics have deficits in constructing imaginary scenes (Hassabis et al 2007).

Spatial memory. Another prominent example of hippocampally-dependent process is spatial memory in rodents (Morris et al 1982) and humans (Eichenbaum et al 1999; Burgess et al 2002).

6.3. The putative role of the hippocampus in other cognitive functions can be accounted for by its role in episodic memory

When we say that the hippocampus is primarily serving a role in storing and retrieving episodic memory, this does not exclude that the hippocampus might be involved in learning or performing other tasks. Cognitive processes interact with each other in complex ways to give rise to an observable behavioral output. That makes it all the more important to dissociate the individual function of each cognitive process and examine how a process interacts with others. In our opinion, some studies mentioned in the previous sections find a hippocampal involvement because the task used in those studies draw on episodic memory for optimal function.

Perception of objects. Where deficits in object perception have been reported in subjects with extensive MTL lesions, it has to be carefully evaluated whether this deficit is due to damage to the hippocampus or to MTL regions outside the hippocampus. For object perception, the mounting evidence is that subjects with damage limited to the hippocampus perform on par with controls, even if patients with broader MTL lesions are impaired (Lee et al 2005; Lee

and Rudebeck 2010; Barense et al 2012). The majority of studies examining the role of the MTL in visual discrimination structures have focused on the perirhinal cortex (Lee et al 2012), since monkeys with lesions of the perirhinal cortex have deficits using conjunctions of visual features (e.g., object perception) to make visual discriminations (Buckley et al 2001; Bussey et al 2002). Importantly, the same monkeys have no deficit in visual discrimination based on simple visual features such as size and color. Furthermore, focal hippocampus lesions in monkeys specifically do not impair object discrimination tasks that are sensitive to lesions of the perirhinal cortex (Saksida et al 2006). Taken together, the results suggest that, if the MTL is involved in object perception at all, it is not due to the involvement of the hippocampus. However, due to methodological concerns some authors suggest that perceptual deficits after damage to the perirhinal cortex might in fact be due to an implicit memory component in the task (Hampton 2005; Suzuki 2009).

Perception of space. The situation is different for the visual perception of spatial scenes, where the hippocampus itself appears to play a role (Lee et al 2005). However, the experimental evidence is far from consistent and the interpretation of these findings remains in dispute. For instance, two patients with focal hippocampal lesions in one study had deficits, while two others, including one with dense amnesia, did not (Hartley et al 2007). The heterogeneity of these findings might be due to a lack of understanding of how subjects solve the tasks. Some subjects might try to rely more than others on their (impaired) long-term memory to solve the task (Suzuki 2009). It is possible, in general, that subjects might perform better in the perceptual task by learning the stimuli across different trials (Shrager et al 2006; Kim et al 2011). Consistent with the hypothesis that perception itself is unaffected in patients with MTL damage, a recent study has found that these patients show the same eye fixation patterns while solving difficult visual discrimination problems (Erez et al 2013). We therefore conclude that the experimental data currently does not show unambiguously that the hippocampus is directly involved in the perception of space.

Language. Since there is no evidence that patients with brain damage restricted to the hippocampus have language deficits immediately following the damage, we conclude that the hippocampus does not have a direct involvement in language. The reduced retention of low-frequency words was observed in patient HM only decades after his surgery, but not immediately afterwards (James and MacKay 2001). Since high-frequency words are unaffected regardless of age (James and MacKay 2001), our interpretation of these results is that low-frequency words are forgotten, if they are not maintained. Maintenance requires either using the words, or covert rehearsal driven by episodic memory, which requires the hippocampus.

Short-term memory. For the Sequence Analysis, the delay between t_0 , at which the experience occurs, and t_1 , at which the memory is retrieved, plays no role. Thus, according to the Sequence Analysis, episodic memories can exist across both short and long retention delays. However, short- and long-term memory cannot be simply distinguished based on the delay between storage and retrieval (Cowan 2008). Short-term memory is thought to differ in essential ways in that it has a limited chunk capacity and to decay over relatively short time scales. When short-term memory is used in a processing system requiring attention, then we refer to this combined system as working memory (Baddeley and Hitch 1974; Cowan 2008). Nonetheless, some authors have suggested that short-term memory is simply activated long-term memory, where the activation has limited capacity and decays in time (Cowan 1995; Ruchkin et al 2003) and that, therefore, short- and long-term memory share the same neural basis (Ranganath and Blumenfeld 2005; Jonides et al 2008). In this case, the working memory deficit in amnesics could be simply explained by their deficit in episodic memory.

However, other authors regard the working memory system as distinct from long-term memory (Baddeley 2012). Even in this case, the working memory deficit in amnesics does not necessarily imply that the hippocampus has a non-episodic-memory function in working memory. At least two alternative explanations are possible. First, findings of working memory deficits might be confounded by extra-hippocampal damage. For instance, even dense amnesics have preserved working memory across short delays – if their intelligence and executive capacities are well-preserved (Baddeley and Wilson 2002), consistent with other findings that in healthy subjects intellectual aptitude is strongly correlated with working memory performance (Daneman and Carpenter 1980; Alloway et al 2009; Alloway and Alloway 2010). Second, working memory deficits were seen either because the tasks involved a long-term-memory component or because the subjects attempted to use a long-term memory strategy to solve the task (Cowan 2008; Baddeley 2012). Finally, we like to emphasize that activation of the medial temporal lobe during active maintenance of novel information (Ranganath and D'Esposito 2001) does not necessarily imply that the hippocampus is directly involved in working memory. Instead, the hippocampus might become active because working memory interacts with long-term memory (Baddeley 2012). Consistent with this interpretation are findings that the strength of the delay activity predicts later long-term memory performance (Ranganath et al 2005).

Temporal associations. Sequence learning has a clear relationship to episodic memory in the Sequence Analysis. In trace conditioning, animals learn to associate an initially neutral stimulus, such as a tone, with a stimulus that elicits an automatic response, such as an electric shock. The crucial point is that the two stimuli do not overlap in time. If they do (delay conditioning), learning is independent of the hippocampus. We suggest that trace conditioning requires the hippocampus because episodic memory is required to learn the task. Since episodes are extended in time in the Sequence Analysis, episodic memory is apt to bridge the temporal gap between the two stimuli (Pyka and Cheng 2014).

Semantic memory. Semantic memories can be formed without a functioning hippocampus, because subjects who became amnesic during childhood can acquire sufficient semantic memory to pass secondary education and reach a normal IQ (Vargha-Khadem et al., 1997). This observation calls into question the hypothesis, such as Klein's (2013) recent suggestion, that a single memory system subserves both semantic and episodic memory. A potential mechanism by which amnesics form new semantic memories is a special learning protocol, called fast mapping, which was recently found to be intact in amnesics (Sharon et al 2011). We therefore conclude that the hippocampus is not a prerequisite for semantic learning. The difficulty that amnesics have in learning new semantic memories (Levy et al 2004), we suggest, are due to the fact that usually episodic memories are used by the neocortex to extract new semantic knowledge (Nadel and Moscovitch 1998; Cheng 2013).

Scene construction. There is currently an unresolved discrepancy between subjects who became amnesic in adulthood and during development. Adulthood amnesics are either impaired at scene construction or use residual hippocampal tissue for scene construction (Mullally et al 2012). However, few developmental amnesics are impaired and, at least, one subject (Jon) does not engage remaining hippocampal tissue during scene construction (Mullally et al 2014). On the other hand, developmental amnesics are equally impaired in their episodic memory performance as compared to adulthood amnesics. These results suggest that scene construction, but not episodic memory, can be accomplished without the hippocampus.

Spatial memory. To account for the strong spatial responses recorded in the rodent hippocampus, O'Keefe and Nadel (1978) suggested that a cognitive map evolved in the hippocampus of non-human mammals to support spatial navigation and that this cognitive

map is used in humans to support episodic memory. The view that spatial information plays a special role for the hippocampus was boosted by the discovery of grid cells in the medial entorhinal cortex, one of the major inputs to the hippocampus (Hafting et al 2005). Alternatively, other authors (Eichenbaum et al 1999; Cheng 2013) have suggested that episodic memory is the primary function of the hippocampus and that spatial information is only one aspect of the content of episodic memory (Tulving 1972), which enters the hippocampus via the grid cells (Cheng and Frank 2011; Azizi et al 2014). Consistent with this view, the other major input to the hippocampus, the lateral entorhinal cortex, does not exhibit spatial coding (Hargreaves et al 2005), and nonspatial information strongly modulates hippocampal spiking activity (Wood et al 1999).

Taken together, the experimental evidence suggests to us that the various functions that have been attributed to the hippocampus can be accounted for by an involvement of episodic memory in the tasks used for testing. Therefore, we conclude that the experimental results are consistent with the hippocampus being dedicated to the storage and retrieval of episodic memories (C2.1).

6.4. The role of other brain regions for episodic memory

Experimental evidence suggests that other brain regions, in particular the prefrontal cortex, are involved in the formation and retrieval of episodic memories. For instance, after lesions of the prefrontal cortex (PFC), patients have a deficit in effortful memory tasks such as recognition, cued-recall and free recall (Wheeler et al 1995). In addition, imaging studies revealed that the PFC is activated during encoding and retrieval of episodic memories (Tulving et al 1994). It was suggested that the two hemispheres are activated asymmetrically during different memory phases with the left more active during encoding and the right more during retrieval (Tulving et al 1994). However, later studies suggest that the left-right asymmetry depends on the content of the memory rather than on the memory phase (Golby et al 2001). The left PFC was more active for verbal tasks, whereas the right PFC was more active in non-verbal tasks. Interestingly, these asymmetries are similar to the verbal/non-verbal asymmetries observed after hippocampal lesions.

However, general episodic memory is only slightly impaired after lesions of the PFC. So patients with restricted frontal lesions are not usually considered amnesic (Wheeler et al 1995). In addition, cognitive deficits are much more widespread after frontal than after hippocampal lesions. The affected functions are collectively referred to as executive control and include, among others, task switching (Milner 1963), decision making (Bechara et al 1994), and working memory (Jacobsen 1935). In summary, to the best of our knowledge, there is no convincing evidence that any other brain region is as central for the formation of episodic memory as the hippocampus. While it is always possible that future studies will reveal such a brain region, until that time, it is most parsimonious to assume that the hippocampus plays a unique role in episodic memory.

The preceding statement does not imply that episodic memory is stored and retrieved in the hippocampus alone. On the contrary, we believe that the hippocampus is part of a network that performs these functions and that the neocortex is critical for processing the sensory information to be stored, for initiating memory retrieval and for processing the retrieved information (Tulving 1995; Nadel and Moscovitch 1998). Specifically, what we mean by "the hippocampus plays a unique role in episodic memory" is that the hippocampus endows the cortico-hippocampal network with a capability that the network does not have without the hippocampus. For instance, a recent modeling study suggested that the hippocampus enables the cortico-hippocampal network to associate two inputs across significant time gaps of 150 ms (Pyka and Cheng 2014). This function emerges from two simple anatomical properties of

the biological network: heterogeneous synaptic conductance delays between neocortex and hippocampus, and a high degree of convergence from cortical to hippocampal cells. We hypothesize that adding more detailed anatomical structures such as intra-hippocampal connections could further increase the time gap across which temporal associations could be formed. Without the hippocampus, the network can still learn associations, but across significantly smaller time gaps of 50ms. Without the neocortex, the model cannot learn any associations. So both neocortex and hippocampus are required, but the hippocampus adds a specialized functionality to the network. We are therefore justified in saying that, in the model, the hippocampus plays a special role in learning an association across larger time gaps. We think that a similar characterization can be applied to the biological cortico-hippocampal network.

In summary, no other brain region is required for episodic memory to the degree that the hippocampus is (C2.2).

7. Neural mechanisms of episodic memory

The representation of episodic memory, on which the Sequence Analysis rests, cannot be studied with purely behavioral readouts. While humans are able to report some information about their internal states verbally, it is doubtful that they have direct access to all the representations in their brain. A case in point is the fact that episodic memory was discovered accidentally as a form of memory only after a certain brain region was removed from patient HM to control his epileptic seizures. To shed light on the neural basis of episodic memory, we therefore have to rely on invasive experiments in nonhuman animals. Since the Sequence Analysis makes no reference to language or subjective experience, it is well suited to study episodic memories in animals.

So far we have argued that mnemonic representations that satisfy the Sequence Analysis form both a maximal and minimal class and that the properties specified in the Sequence Analysis serve inductive and explanatory purposes (section 5). What remains to be done to conclude that episodic memories as explicated in the Sequence Analysis form a natural kind is to show that episodic memories share these properties because of some uniform causal mechanisms. We have presented evidence in section 6 that the hippocampus is the principle anatomical substrate of episodic memory. In this section, we change the level of description and zoom in into the neurophysiological mechanisms within the hippocampus (C3). We argue that neural processes in the hippocampus provide uniform causal mechanisms for the conditions that the Sequence Analysis places on encoding (S5 and S6), storage (S1 and S2) and retrieval (S7). We may presuppose in the following that there are neural mechanisms that ensure veridical perception (S3 and S4).

7.1. Compression of experiential sequences and mnemonic sequences in the hippocampus

Principal neurons in the hippocampus are active in specific, CIRCUMSCRIBED spatial regions (place fields). O'Keefe and Dostrovsky (1971) therefore called these neurons place cells (Fig. 1a). While it remains unresolved whether spatial information plays a special role in episodic memory (we discussed this point in 6.3), here we only rely on the uncontroversial observation that place cells are activated at certain locations in the environment. We suggest that the activations of place cells p_1, p_2, p_3 signal the occurrences of events e_1^*, e_2^*, e_3^* that occur at the location of their respective place fields. If the place cells are sorted according to their place field locations, then the place cells fire in a temporal sequence that could represent the experience of an episode $E^* = \langle e_1^*, e_2^*, e_3^* \rangle$ when the animal runs along a trajectory (S3).

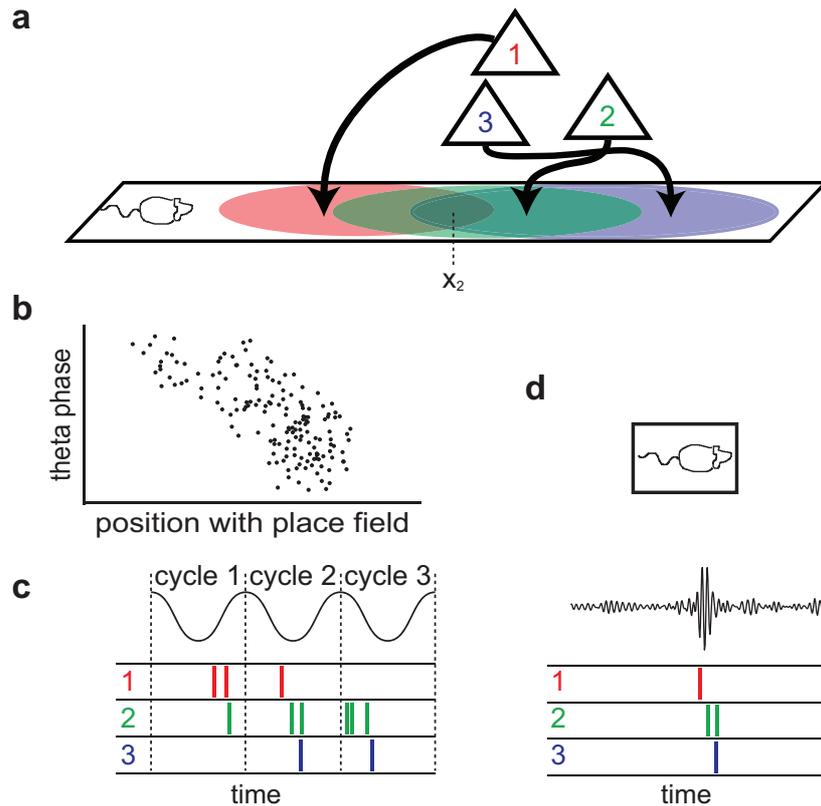


Figure 1: *Schematic illustration of neural activity of hippocampal neurons.* **a:** As the animal explores the linear track, place cells (1,2,3) fire spikes when the animal is located in a circumscribed region in space, the place field (indicated by three colored ellipses). The location marked by x_2 is used in an example discussed in the main text. **b:** In addition, the spiking of place cells is modulated by the phase of the theta oscillation. Each dot marks the theta phase and position of the animal when a neuron fired a spike as the animal run from left to right. Early in the place field, spikes occur at late phases. Just before the animal exits the place field, spikes occur at early theta phases. The relationship between theta phase and animal position is known as theta phase precession. **c:** When spiking of a group of place cells is analyzed within one cycle of the theta oscillation (black trace at the top), temporal sequences emerge across neurons (theta sequences). Dotted lines illustrate the (arbitrary) beginning and end of theta cycles. Cycle 2 occurs roughly when the animal is located in position x_2 . **d:** During the offline state, sharp wave/ ripples occur in the local field potential (black trace in middle, filtered between 150 and 250 Hz) and place cells are reactivated in a sequence that is related to the theta sequences.

According to the Sequence Analysis, to form episodic memory in the hippocampus, a mnemonic representation of E has to be stored in the hippocampus, where the relation of E to E^* satisfies condition (S5). Many authors have suggested that the hippocampal circuitry is optimized for storing neural sequences. There is widespread agreement in neuroscience that mnemonic representations, i.e., the neural substrate of memories, are stored in the weights of the synaptic connections between neurons. More specifically, it has been suggested that the dense recurrent network in subarea CA3 is well suited to generate neural sequences (Levy 1996; Amarasingham and Levy 1998; Wallenstein et al 1998; Lisman 1999; Azizi et al 2013; Cheng 2013). To store a memory, the experience must drive some appropriate change in the synaptic weights. These changes are referred to as synaptic plasticity and require precise timing relationships between the spikes fired by the connected neurons. The generation of a

mnemonic representation is not trivial since there is a mismatch of timescales: The experienced episode E^* unfolds over seconds (the time the animal requires to run through the place fields of p_1, p_2, p_3), whereas spike-timing dependent plasticity requires spikes to co-occur within tens of milliseconds. This mismatch of timescales raises the question of how the experience of E^* could causally ground the mnemonic representation of E as required by condition (S6). The answer could be provided by a compression mechanism that is based on theta phase precession (O'Keefe and Recce 1993; Dragoi and Buzsáki 2006). This compression mechanism generates a representation of the behavioral sequence at the shorter timescale required for synaptic plasticity. We explain how this mechanism works in more detail in the next paragraph, but a detailed understanding of phase precession is not required to follow the subsequent arguments.

Theta oscillations (5-12 Hz) occur when rats are actively involved in a task, which we refer to as online state (Buzsaki 1989). During spatial exploration, place cells initially fire spikes at the peak of this theta oscillation and then at earlier and earlier phases of the theta oscillation as the animal traverses the place field (from left to right in Fig. 1b). The relationship between the location within the place field and the theta phase, when spikes are fired, is called theta phase precession. Due to phase precession, place cells with overlapping, but non-identical, place fields fire spikes at different phases of theta. For instance, assume the animal is located at x_2 and the theta oscillation is in cycle 2 (Fig. 1a,c). Then cell p_1 fires spikes early in theta cycle 2 because x_2 is near the exit portion of p_1 's place field. The same location x_2 falls into the middle of p_2 's place field, and so p_2 fires spikes in the middle of theta cycle 2. Finally, p_3 fires spikes late in theta cycle 2 since x_2 lies in the entry portion of p_3 's place field. As a result, the spikes of the three cells are temporally ordered within the single theta cycle 2 (Fig. 1c). Furthermore, the temporal sequence of the spikes of p_1, p_2, p_3 corresponds to the spatial succession of their respective place fields x_1, x_2, x_3 (Skaggs et al 1996). Since the temporal sequence is played out within a single theta cycle, the spikes occur within tens of milliseconds, on a much shorter time scale than the behavioral sequence (Dragoi and Buzsáki 2006) and well within the time window of synaptic plasticity.

While brain recordings in humans are rare, one recent experiment found that neuronal activity in the hippocampus during memory encoding is sequentially organized (Paz et al 2010), suggesting that the observations of neural sequences in rodent hippocampus might generalize to humans.

7.2. *Offline sequential activity and replay in the hippocampus*

In the offline state, i.e., when the animal sits quietly or is asleep, sharp-wave ripples (SWRs, Fig. 1D) dominate network oscillations. SWRs have been observed in rodents (Buzsáki et al 1983), non-human primates (Skaggs et al 2007) and the human hippocampus and entorhinal cortex (Bragin et al 1999), suggesting that SWRs are part of a general, conserved mechanism. Concurrent with SWRs in the hippocampus, populations of place cells fire spikes in a temporal sequence within a 50-400 ms time-window (Lee and Wilson 2004).

The critical point is that the sequence of spiking, e.g., of p_1, p_2, p_3 , in the offline state is correlated with and influenced by preceding online activity (Fig. 1d; for a review, see Buhry et al 2011). Pavlides and Winson (1989) were the first to report that individual place cells that were active during behavior were more likely to be active again during subsequent sleep and quiescence than those place cells that were not active during explorations. Subsequent studies reported the reactivation of pairs of cells (Wilson and McNaughton 1994; Kudrimoti et al 1999), which also preserve their ordering (Skaggs and McNaughton 1996). Most importantly, populations of neurons in the offline state fire in a sequence that correlates with the sequence, in which they were active at an earlier time in the online state (Nádasy et al 1999; Lee and

Wilson 2002). Thus, sequential neural activity in the offline state is a replay of sequential activity in prior experience. Replay has been observed across species and brain regions, such as rodent hippocampus (Lee and Wilson 2002); rodent PFC (Euston et al 2007); primate motor, somatosensory, and parietal cortex (but not prefrontal cortex) (Hoffman and McNaughton 2002); and during free recall of movie sequences in humans (Gelbard-Sagiv et al 2008). This is consistent with the mnemonic content E being grounded ontologically in the experiential base E^* (S5). Together with the results discussed in the preceding section, this means that the sequential representation of mnemonic content in the offline state is aligned with the sequential representation of the experiential base in the online state (C3.1).

As discussed in 3.4, in the Sequence Analysis we distinguish between an actual and temporally enduring mnemonic representation (S2) and a possible and only instantaneous mnemonic simulation (S7). We have recently shown in a computational model that the weights in a neural network can be set up such that the network can spontaneously generate neuronal sequences (Azizi et al 2013). In other words, the enduring structure of the network, i.e., the matrix of synaptic weights, can represent sequential content. This enduring representation of a sequence can be read out resulting in a replay of the sequence and thus generating an instantaneous simulation of the sequence.

Corballis (2013) has argued recently that neural sequences are an indication that non-human animals have episodic memory. However, his view of episodic memory depends on the mental time travel idea (Suddendorf and Corballis 1997) and he needs to argue that neural sequences are a correlate of the subjective experience of the animal, which Suddendorf has rejected (Suddendorf 2013). Condition S7 in the Sequence Analysis remains agnostic about the subjective experience during the mnemonic simulation.

Experimental evidence suggests that the hippocampus can combine segments of previously experienced sequences into a new sequence that was never experienced as a whole (Gupta et al 2010). In addition, our computational simulations indicate that a network can generate a variety of sequences and the same neurons can participate in different sequence representations (Azizi et al 2013). In summary, we can conclude that experimental and modeling results suggest that the hippocampus provides a uniform causal mechanism for the compositional representation and mnemonic simulation of experientially grounded episodes (C3.2).

7.3. Replay in the hippocampus is linked to the formation and consolidation of episodic memory

The final crucial aspect missing from our discussion is evidence for a memory trace that causally links the experience of the episode to the mnemonic representation (S6). The most relevant studies in this regard are those that examine the link between offline sequences and the systems consolidation process. Much time and effort has been devoted to understanding the exact properties and neural mechanisms of consolidation. Buzsaki (1989) proposed that, first, a labile memory trace is formed in the hippocampus during the online state. Then, during subsequent offline states, hippocampal replay gradually transfers the memory trace to neocortical areas (Buzsaki 1989; McClelland et al 1995).

To examine the functional role of neural sequences, a number of studies exploited its co-occurrence with SWR (Fig. 1d). Mounting experimental evidence suggests that SWRs are important for learning and memory. For instance, the rate of SWRs was found to be higher in a novel than in a familiar part of an environment and so is the spiking probability of place cells (Cheng and Frank 2008). SWRs were observed to increase during slow-wave sleep after learning (Eschenko et al., 2008; Ramadan et al., 2009). The number of rhinal SWRs in humans

during a daytime nap appears to be correlated with the number of successfully recalled items learned prior to sleep (Axmacher et al 2008). Disrupting SWR in rats during sleep after a learning session interferes with the formation of long-term memories (Girardeau et al 2009). Disrupting SWRs in rat hippocampus during the awake state disrupted learning a spatial working memory task (Jadhav et al., 2012). Taken together, these results suggest that offline replay of sequences in the hippocampus are involved in maintaining the memory trace of episodic memories (Cheng and Werning 2013), and that the memory trace causally links experiences to mnemonic representation (C3.3).

For completeness, we note that not all sequential activity in the hippocampus is causally grounded in previously experienced sequences. During exploration, theta sequences appear to begin in the past and sweep to anticipated locations (Gupta et al 2012). These results indicate that memories are retrieved during exploration, as would be required if memory influences future behaviors. The temporal order of offline neural sequences also does not strictly correlate with the temporal order of online neural sequences that have occurred in the past. In the awake state, Gupta et al. (2010) observed in the hippocampus offline neural sequences that corresponded to trajectories that the animal had never traveled. Dragoi and Tonegawa (2011) reported evidence for pre-play. Neural sequences recorded during rest were predictive of the sequence of the neurons' place fields on a linear track that the animal had never experienced before. These results suggest that the hippocampal network generates spontaneous sequences that are constrained by its network architecture and dynamics (Buhry et al 2011; Azizi et al 2013). The fact that sequential representations in the hippocampus may play a role also in the exploration of future possible trajectories hints at a possible relationship between episodic memory and anticipating the future.

8. Conclusion

The starting point for us in this paper has been the insight that the two questions: “What is episodic memory?” and “Is episodic memory a natural kind?” are inherently connected to each other. The first question cannot be answered meaningfully without aiming at a positive answer to the second: It is one thing to have a notion of episodic memory that, for better or worse, matches our conceptual intuitions, our linguistic practice and perhaps some introspective phenomenology. However to enable induction and explanation in science, one should make sure that one is referring to a natural kind when speaking of episodic memory. For the purposes of sciences such as psychology and neuroscience, identifying natural kinds with homeostatic property clusters, as we did, seems most fruitful.

In turn, an answer to the second question has to be assessed in light of the consequences it has for the first. It would not suffice to enlist a number of neural mechanisms that amount to particular psychological properties and label them “episodic memory”. What has to be done in addition is to show that uniform causal mechanisms explain why the psychological properties are shared such that the cluster of those properties subserve inductive and explanatory purposes of what we are to understand is episodic memory.

In search for an answer to the conditional question “What is episodic memory if it is a natural kind?” we have tied analytical and empirical approaches most closely together. This has made our paper a truly interdisciplinary one, a combination of philosophy and neuroscience. In the Sequence Analysis episodic memory is conceived of as a factive, knowledge-like state that consists of an experientially based mnemonic representation and has the potential for mnemonic simulation. We have stressed the sequential character of the mnemonic content as being an episode. That is, a temporally ordered list of concrete and particular events.

We have tried to substantiate the Sequence Analysis of episodic memory as corresponding to a natural kind by proceeding along three empirical cornerstones: Is there psychological evidence that a violation of any of the conditions of the Sequence Analysis amounts to a deficiency in episodic memory and is it assured that no form of memory or cognitive process but episodic memory fulfills them? Even though we could not exhaust the empirical literature here, our answer was affirmative. Do the empirical data support a claim of what the principal anatomical substrate of episodic memory is, given that the Sequence Analysis holds? We have pointed to a great deal of evidence that there is one: the hippocampus. Finally, do we know the neural activities in the hippocampus onto which we can pin down causal mechanisms in order to explain the psychological states and processes appealed to by the Sequence Analysis? Also here we could call on a body of evidence from neuroscience.

We regard the mutually supporting interaction between philosophical analysis and neuroscientific evidence as a particular strength of our approach. This is something that the alternative accounts of episodic memory mentioned earlier in this article have yet to deliver. Furthermore, our notion of episodic memory does not depend on a classical taxonomical classification and is therefore independent of whether or not the contrasting notion of semantic memory or the superordinate notion of declarative memory correspond to natural kinds.

Of course, our conclusions depend on a particular selection and interpretation of experimental results. We admit that it is likely that details, or even substantial aspects, of our interpretations have to be modified in the future. Whether it is possible to adjust the framework to stay consistent with new findings has to be seen, when the time comes. The utility of our account is that it provides a uniform and parsimonious framework for the interpretation of a highly diverse set of experimental results. We hope that this framework will drive the vibrant research on episodic memory forward.

9. Appendix

In section 3.4 we distinguish between an actual and temporally enduring mnemonic representation and a possible and only instantaneous temporally explicit mnemonic simulation. The mnemonic representation is a compositional representation in which the temporal succession of the events in the episode is encoded by some structure among the representational constituents. This can be formally accounted for in the following way:

- (R1) Let $\mathbb{R} = (R, \{s\})$ be a representational structure, where R is the set of representational states and $s: R \times R \rightarrow R$ a thereon defined binary and recursive structure building operation; if $r'' = s(r, r')$, r and r' are called the representational constituents of r'' .
- (R2) Let $\mathbb{V} = (V, \{\langle \cdot, \cdot \rangle\})$ be an event structure where V is the set of events (including episodes) and $\langle \cdot, \cdot \rangle$ the previously defined binary and recursive episode building operation $\langle \cdot, \cdot \rangle: V \times V \rightarrow V$.
- (R3) Let $\mu: R \rightarrow V$ be a mapping from the set of representational states into the set of events. $\mu(r) = e$ means that e is the event or episode represented by the primitive or complex representational state r .
- (R4) Then the mapping μ is a homomorphism from the representational structure \mathbb{R} into the event structure \mathbb{V} . That is, for every $r, r' \in R$, $\mu(s(r, r')) = \langle \mu(r), \mu(r') \rangle$.

Statement (R4) is equivalent to the claim that the representational structure is compositional: The content μ of a complex representational state $s(r, r')$ is a structure-dependent function, namely $\langle \cdot, \cdot \rangle$, of the contents of its representational constituents r and r' , i.e., $\mu(r)$ and $\mu(r')$.

In a temporally explicit mnemonic simulation the temporal succession of events in the object domain is represented itself by a temporal succession of events in the representational domain. We can formally account for this in the following way:

- (Q1) Let $\mathbb{V} = (V, \{\langle \cdot, \cdot \rangle\})$ again be our event structure, where V is the set of events and $\langle \cdot, \cdot \rangle$ the episode building operation.
- (Q2) Let $\mathbb{Q} = (Q, \{\langle \cdot, \cdot \rangle\})$ be the simulational structure, which itself contains a set of events Q , i.e., $Q \subseteq V$, together with the operation $\langle \cdot, \cdot \rangle$, which builds temporally succeeding sequences of events. \mathbb{Q} hence is a substructure of \mathbb{V} .
- (Q3) Let the mapping $v: Q \rightarrow V$ assign each primitive or complex simulational state the primitive or complex event it, by simulation, represents.
- (Q4) Then the mapping v satisfies the following condition: For every $q' \in Q$, $v(\langle q, q' \rangle) = \langle v(q), v(q') \rangle$.

In this account, the representational content of a temporal succession of simulational states is just the episode comprising the temporally succeeding representational contents of each simulational state. Mathematically speaking, v , unlike μ , is not only a homomorphism but an endomorphism.

References

- Aggleton JP, Shaw C, Gaffan EA (1992) The performance of postencephalitic amnesic subjects on two behavioural tests of memory: concurrent discrimination learning and delayed matching-to-sample. *Cortex* 28:359–72.
- Agster KL, Fortin NJ, Eichenbaum H (2002) The hippocampus and disambiguation of overlapping sequences. *J Neurosci* 22:5760–8.
- Alloway TP, Alloway RG (2010) Investigating the predictive roles of working memory and IQ in academic attainment. *J Exp Child Psychol* 106:20–9. doi: 10.1016/j.jecp.2009.11.003
- Alloway TP, Gathercole SE, Kirkwood H, Elliott J (2009) The cognitive and behavioral characteristics of children with low working memory. *Child Dev* 80:606–21. doi: 10.1111/j.1467-8624.2009.01282.x
- Aly M, Ranganath C, Yonelinas AP (2013) Detecting changes in scenes: the hippocampus is critical for strength-based perception. *Neuron* 78:1127–37. doi: 10.1016/j.neuron.2013.04.018
- Amarasingham A, Levy WB (1998) Predicting the distribution of synaptic strengths and cell firing correlations in a self-organizing, sequence prediction model. *Neural Comput* 10:25–57.
- Axmacher N, Elger CE, Fell J (2008) Ripples in the medial temporal lobe are relevant for human memory consolidation. *Brain* 131:1806–17. doi: 10.1093/brain/awn103
- Azizi AH, Schieferstein N, Cheng S (2014) The transformation from grid cells to place cells is robust to noise in the grid pattern. *Hippocampus* 24:912–9. doi: 10.1002/hipo.22306

- Azizi AH, Wiskott L, Cheng S (2013) A computational model for preplay in the hippocampus. *Front Comput Neurosci* 7:161. doi: 10.3389/fncom.2013.00161
- Babb SJ, Crystal JD (2005) Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learn Motiv* 36:177–189. doi: 10.1016/j.lmot.2005.02.009
- Baddeley A (2012) Working memory: theories, models, and controversies. *Annu Rev Psychol* 63:1–29. doi: 10.1146/annurev-psych-120710-100422
- Baddeley A, Wilson BA (2002) Prose recall and amnesia: implications for the structure of working memory. *Neuropsychologia* 40:1737–43.
- Baddeley AD, Hitch G (1974) Working Memory. *Psychol Learn Motiv* 8:47–89. doi: 10.1016/S0079-7421(08)60452-1
- Barens MD, Groen II a, Lee ACH, et al (2012) Intact memory for irrelevant information impairs perception in amnesia. *Neuron* 75:157–67. doi: 10.1016/j.neuron.2012.05.014
- Bauer PJ, Doydum AO, Pathman T, et al (2012) It’s all about location, location, location: children's memory for the “where” of personally experienced events. *J Exp Child Psychol* 113:510–22. doi: 10.1016/j.jecp.2012.06.007
- Baxter MG (2009) Involvement of Medial Temporal Lobe Structures in Memory and Perception. *Neuron* 61:667–677. doi: 10.1016/j.neuron.2009.02.007
- Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50:7–15. doi: 10.1016/0010-0277(94)90018-3
- Bedford FL (1997) False categories in cognition: the Not-the-Liver fallacy. *Cognition* 64:231–48.
- Bermudez JL (1995) Nonconceptual Content: From Perceptual Experience to Subpersonal Computational States. *Mind Lang* 10:333–369. doi: 10.1111/j.1468-0017.1995.tb00019.x
- Bernecker S (2010) *Memory : a philosophical study*. Oxford University Press
- Bernecker S (2008) *The Metaphysics of Memory*. New York, NY
- Berry CJ, Shanks DR, Henson RN a (2008) A single-system account of the relationship between priming, recognition, and fluency. *J Exp Psychol Learn Mem Cogn* 34:97–111. doi: 10.1037/0278-7393.34.1.97
- Bieri P (1995) Why is consciousness puzzling? In: Metzinger T (ed) *Conscious Exp*. Schoening, Imprint Academic, Paderborn, pp 45–60
- Bischof N (1980) On the Pyhlogeny of Human Morality. In: Stent GS (ed) *Moral. as a Biol. phenomenon*. Rep. Dahlem Work. Biol. Morals. Universtity of California Press, Berkeley, pp 48–66

- Bischof-Köhler D (1985) Zur Phylogenese menschlicher Motivation [On the phylogeny of human motivation]. In: Eckensberger LH, Baltes MM (eds) *Emot. und Reflexivität*. Urban & Schwarzenberg, München, pp 3–47
- Bontempi B, Laurent-Demir C, Destrade C, Jaffard R (1999) Time-dependent reorganization of brain circuitry underlying long-term memory storage. *Nature* 400:671–5. doi: 10.1038/23270
- Boyd R (1991) Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philos Stud* 61:127–148. doi: doi: 10.1007/bf00385837
- Boyd R (1999) Kinds, Complexity and Multiple realization. *Philos Stud* 95:67–98. doi: doi: 10.1023/a%253a1004511407133
- Bragin A, Engel J, Wilson C, et al (1999) High-frequency oscillations in human brain. *Hippocampus* 9:137–142. doi: 10.1002/(SICI)1098-1063(1999)9:2%3C137::AID-HIPO5%3E3.0.CO;2-0
- Buckley MJ, Booth MC, Rolls ET, Gaffan D (2001) Selective perceptual impairments after perirhinal cortex ablation. *J Neurosci* 21:9824–36.
- Buhry L, Azizi AH, Cheng S (2011) Reactivation, replay, and preplay: How it might all fit together. *Neural Plast* 2011:1–11. doi: 10.1155/2011/203462
- Burgess N, Maguire E a, O’Keefe J (2002) The human hippocampus and spatial and episodic memory. *Neuron* 35:625–41. doi: 10.1016/S0896-6273(02)00830-9
- Bussey TJ, Saksida LM, Murray EA (2002) Perirhinal cortex resolves feature ambiguity in complex visual discriminations. *Eur J Neurosci* 15:365–74.
- Buzsáki G (1989) Two-stage model of memory trace formation: A role for “noisy” brain states. *Neuroscience* 31:551–570.
- Buzsáki G, Leung LW, Vanderwolf CH (1983) Cellular bases of hippocampal EEG in the behaving rat. *Brain Res* 287:139–71.
- Cheng S (2013) The CRISP theory of hippocampal function in episodic memory. *Front Neural Circuits* 7:88. doi: 10.3389/fncir.2013.00088
- Cheng S, Frank L (2011) The structure of networks that produce the transformation from grid cells to place cells. *Neuroscience* 197:293–306. doi: 10.1016/j.neuroscience.2011.09.002
- Cheng S, Frank LM (2008) New experiences enhance coordinated neural activity in the hippocampus. *Neuron* 57:303–13. doi: 10.1016/j.neuron.2007.11.035
- Cheng S, Werning M (2013) Composition and replay of mnemonic sequences : The contributions of REM and slow-wave sleep to episodic memory. *Behav Brain Sci* 36:610–611. doi: 10.1017/S0140525X13001234
- Clayton NS, Bussey TJ, Dickinson A (2003) Can animals recall the past and plan for the future? *Nat Rev Neurosci* 4:685–91. doi: 10.1038/nrn1180

- Clayton NS, Dickinson A (1998) Episodic-like memory during cache recovery by scrub jays. *Nature* 395:272–4. doi: 10.1038/26216
- Cohen J, Meskin A (2004) On the Epistemic Value of Photographs. *J Aesthet Art Crit* 62:197–210.
- Conway M a (2009) Episodic memories. *Neuropsychologia* 47:2305–13. doi: 10.1016/j.neuropsychologia.2009.02.003
- Corballis MC (2013) Mental time travel: a case for evolutionary continuity. *Trends Cogn Sci* 17:5–6. doi: 10.1016/j.tics.2012.10.009
- Cowan N (2008) What are the differences between long-term, short-term, and working memory? *Prog Brain Res* 169:323–38. doi: 10.1016/S0079-6123(07)00020-9
- Cowan N (1995) *Attention and Memory*. Oxford University Press
- Daneman M, Carpenter PA (1980) Individual differences in working memory and reading. *J Verbal Learning Verbal Behav* 19:450–466. doi: 10.1016/S0022-5371(80)90312-6
- Davidson D (1980) [originally published in 1967] *The Logical Form of Action Sentences. Essays on Actions and Events*. Clarendon Press, Oxford, pp 105–121
- Dere E, Huston JP, De Souza Silva M a (2005) Episodic-like memory in mice: simultaneous assessment of object, place and temporal order memory. *Brain Res Brain Res Protoc* 16:10–9. doi: 10.1016/j.brainresprot.2005.08.001
- Dew ITZ, Cabeza R (2011) The porous boundaries between explicit and implicit memory: behavioral and neural evidence. *Ann N Y Acad Sci* 1224:174–90. doi: 10.1111/j.1749-6632.2010.05946.x
- Dragoi G, Buzsáki G (2006) Temporal encoding of place sequences by hippocampal cell assemblies. *Neuron* 50:145–157. doi: 10.1016/j.neuron.2006.02.023
- Dragoi G, Tonegawa S (2011) Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature* 469:397–401. doi: 10.1038/nature09633
- Eichenbaum H, Dudchenko P, Wood E, et al (1999) The hippocampus, memory, and place cells: is it spatial memory or a memory space? *Neuron* 23:209–26.
- Erez J, Lee ACH, Barense MD (2013) It does not look odd to me: perceptual impairments and eye movements in amnesic patients with medial temporal lobe damage. *Neuropsychologia* 51:168–80. doi: 10.1016/j.neuropsychologia.2012.11.003
- Euston DR, Tatsuno M, McNaughton BL (2007) Fast-Forward Playback of Recent Memory Sequences in Prefrontal Cortex During Sleep. *Science (80-)* 318:1147–1150. doi: 10.1126/science.1148979
- Ezzyat Y, Davachi L (2011) What constitutes an episode in episodic memory? *Psychol Sci* 22:243–52. doi: 10.1177/0956797610393742

- Fortin NJ, Agster KL, Eichenbaum HB (2002) Critical role of the hippocampus in memory for sequences of events. *Nat Neurosci* 5:458–62. doi: 10.1038/nn834
- Friedman WJ (1993) Memory for the time of past events. *Psychol Bull* 113:44–66. doi: 10.1037/0033-2909.113.1.44
- Gelbard-Sagiv H, Mukamel R, Harel M, et al (2008) Internally Generated Reactivation of Single Neurons in Human Hippocampus During Free Recall. *Science* (80-) 322:96–101. doi: 10.1126/science.1164685
- Girardeau G, Benchenane K, Wiener SI, et al (2009) Selective suppression of hippocampal ripples impairs spatial memory. *Nat Neurosci* 12:1222–1223. doi: 10.1038/nn.2384
- Golby AJ, Poldrack RA, Brewer JB, et al (2001) Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain* 124:1841–54. doi: 10.1093/brain/124.9.1841
- Goldman AI (1986) *Epistemology and Cognition*. Harvard University Press, Cambridge, MA
- Graf P, Schacter DL (1985) Implicit and explicit memory for new associations in normal and amnesic subjects. *J Exp Psychol Learn Mem Cogn* 11:501–18.
- Grush R (2004) The emulation theory of representation: Motor control, imagery, and perception. *Behav Brain Sci*. doi: 10.1017/S0140525X04000093
- Gupta AS, van der Meer MA, Touretzky DS, Redish DD (2012) Segmentation of spatial experience by hippocampal θ sequences. *Nat Neurosci* 15:1032–1039. doi: 10.1038/nn.3138
- Gupta AS, van der Meer MAA, Touretzky DS, Redish AD (2010) Hippocampal replay is not a simple function of experience. *Neuron* 65:695–705. doi: 10.1016/j.neuron.2010.01.034
- Hafting T, Fyhn M, Molden S, et al (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436:801–806. doi: 10.1038/nature03721
- Hampton RR (2005) Monkey perirhinal cortex is critical for visual memory, but not for visual perception: reexamination of the behavioural evidence from monkeys. *Q J Exp Psychol B* 58:283–99. doi: 10.1080/02724990444000195
- Hannula DE, Ranganath C (2009) The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron* 63:592–9. doi: 10.1016/j.neuron.2009.08.025
- Harand C, Bertran F, La Joie R, et al (2012) The hippocampus remains activated over the long term for the retrieval of truly episodic memories. *PLoS One* 7:e43495. doi: 10.1371/journal.pone.0043495
- Hargreaves EL, Rao G, Lee I, Knierim JJ (2005) Major dissociation between medial and lateral entorhinal input to dorsal hippocampus. *Science* 308:1792–4. doi: 10.1126/science.1110449

- Hartley T, Bird CM, Chan D, et al (2007) The hippocampus is required for short-term topographical memory in humans. *Hippocampus* 17:34–48. doi: 10.1002/hipo.20240
- Hassabis D, Kumaran D, Vann SD, Maguire EA (2007) Patients with hippocampal amnesia cannot imagine new experiences. *Proc Natl Acad Sci U S A* 104:1726–1731. doi: 10.1073/pnas.0610561104
- Hasselmo ME (2012) *How We Remember: Brain Mechanisms of Episodic Memory*. MIT Press, Cambridge, MA
- Henke K (2010) A model for memory systems based on processing modes rather than consciousness. *Nat Rev Neurosci* 11:523–32. doi: 10.1038/nrn2850
- Hodges W (2001) Formal Features of Compositionality. *J Logic, Lang Inf* 10:7–28.
- Hoffman KL, McNaughton BL (2002) Coordinated Reactivation of Distributed Memory Traces in Primate Neocortex. *Science* (80-) 297:2070–2073. doi: 10.1126/science.1073538
- Holdstock JS, Gaffan D, Bussey TJ, Bussey TJ, Bussey TJ, Bussey TJ (2000) Perceptual and mnemonic matching-to-sample in humans: contributions of the hippocampus, perirhinal and other medial temporal lobe cortices. *Cortex* 36:301–22.
- Holdstock JS, Shaw C, Aggleton JP (1995) The performance of amnesic subjects on tests of delayed matching-to-sample and delayed matching-to-position. *Neuropsychologia* 33:1583–96.
- Jacobsen CF (1935) Functions of frontal association areas in primates. *Arch Neurol Psychiatry* 33:558. doi: 10.1001/archneurpsyc.1935.02250150108009
- James LE, MacKay DG (2001) H.M., word knowledge, and aging: support for a new theory of long-term retrograde amnesia. *Psychol Sci* 12:485–92.
- Jonides J, Lewis RL, Nee DE, et al (2008) The mind and brain of short-term memory. *Annu Rev Psychol* 59:193–224. doi: 10.1146/annurev.psych.59.103006.093615
- Kahana MJ, Howard MW, Polyn SM (2008) Associative Retrieval Processes in Episodic Memory. In: Byrne JH (ed) *Learn. Mem. A Compr. Ref.* Academic Press, Oxford, pp 467–490
- Kelly SD (2001) The Non-conceptual Content of Perceptual Experience: Situation Dependence and Fineness of Grain. *Philos Phenomenol Res* 62:601–608. doi: 10.1111/j.1933-1592.2001.tb00076.x
- Kim S, Jeneson A, van der Horst AS, et al (2011) Memory, visual discrimination performance, and the human hippocampus. *J Neurosci* 31:2624–9. doi: 10.1523/JNEUROSCI.5954-10.2011
- Klein SB (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* 7:3. doi: 10.3389/fnbeh.2013.00003

- Köhler W (1925) *The mentality of apes*. Routledge, Trench, Trubner & Co., Ltd., London
- Kudrimoti HS, Barnes CA, McNaughton BL (1999) Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. *J Neurosci* 19:4090–101.
- Lee ACH, Bussey TJ, Murray E a, et al (2005) Perceptual deficits in amnesia: challenging the medial temporal lobe “mnemonic” view. *Neuropsychologia* 43:1–11. doi: 10.1016/j.neuropsychologia.2004.07.017
- Lee ACH, Rudebeck SR (2010) Human medial temporal lobe damage can disrupt the perception of single objects. *J Neurosci* 30:6588–94. doi: 10.1523/JNEUROSCI.0116-10.2010
- Lee ACH, Yeung L-K, Barense MD (2012) The hippocampus and visual perception. *Front Hum Neurosci* 6:91. doi: 10.3389/fnhum.2012.00091
- Lee AK, Wilson M a (2004) A combinatorial method for analyzing sequential firing patterns involving an arbitrary number of neurons based on relative time order. *J Neurophysiol* 92:2555–2573. doi: 10.1152/jn.01030.2003
- Lee AK, Wilson MA (2002) Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron* 36:1183–1194.
- Levy DA, Bayley PJ, Squire LR (2004) The anatomy of semantic knowledge: medial vs. lateral temporal lobe. *Proc Natl Acad Sci U S A* 101:6710–5. doi: 10.1073/pnas.0401679101
- Levy WBW (1996) A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus* 6:579–590. doi: 10.1002/(SICI)1098-1063(1996)6:6<579::AID-HIPO3>3.0.CO;2-C
- Lewis D (1973) *Counterfactuals*. Blackwell, Oxford
- Lindsay DS, Allen BP, Chan JCK, Dahl LC (2004) Eyewitness suggestibility and source similarity: Intrusions of details from one event into memory reports of another event. *J Mem Lang* 50:96–111. doi: 10.1016/j.jml.2003.08.007
- Lisman JE (1999) Relating hippocampal circuitry to function: recall of memory sequences by reciprocal dentate-CA3 interactions. *Neuron* 22:233–42.
- Loftus EF (1993) The reality of repressed memories. *Am Psychol* 48:518–37.
- Loftus EF, Palmer JC (1974) Reconstruction of automobile destruction: An example of the interaction between language and memory. *J Verbal Learning Verbal Behav* 13:585–589. doi: 10.1016/S0022-5371(74)80011-3
- Loftus EF, Pickrell JE (1995) The formation of false memories. *Psychiatr Ann* 25:720–725.
- Machery E (2009) *Doing Without Concepts*. Oxford University Press, Oxford

- Marsh EJ, Eslick AN, Fazio LK (2008) False Memories. In: Byrne JH (ed) *Learn. Mem. A Compr. Ref.* Academic Press, Oxford, pp 221–238
- Martin CB, Deutscher M (1966) Remembering. *Philos Rev* 75:161. doi: 10.2307/2183082
- Martin-Ordas G, Haun D, Colmenares F, Call J (2010) Keeping track of time: evidence for episodic-like memory in great apes. *Anim Cogn* 13:331–40. doi: 10.1007/s10071-009-0282-4
- McCabe DP, Roediger HL, Karpicke JD (2011) Automatic processing influences free recall: converging evidence from the process dissociation procedure and remember-know judgments. *Mem Cognit* 39:389–402. doi: 10.3758/s13421-010-0040-5
- McClelland JLL, McNaughton BLB, O'Reilly RRC (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.
- McKoon G, Ratcliff R (1986) Automatic activation of episodic information in a semantic memory task. *J Exp Psychol Learn Mem Cogn* 12:108–15.
- Michaelian K (2010) Is memory a natural kind? *Mem Stud* 4:170–189. doi: 10.1177/1750698010374287
- Michaelian K (2011a) Generative memory. *Philos Psychol* 24:323–342. doi: 10.1080/09515089.2011.559623
- Michaelian K (2011b) The information effect: constructive memory, testimony, and epistemic luck. *Synthese* 190:2429–2456. doi: 10.1007/s11229-011-9992-7
- Milner B (1963) Effects of Different Brain Lesions on Card Sorting. *Arch Neurol* 9:90. doi: 10.1001/archneur.1963.00460070100010
- Morris R, Garrud P, Rawlins J, O'Keefe J (1982) Place navigation impaired in rats with hippocampal lesions. *Nature* 297:681–683.
- Morris RG (2001) Episodic-like memory in animals: psychological criteria, neural mechanisms and the value of episodic-like tasks to investigate animal models of neurodegenerative disease. *Philos Trans R Soc Lond B Biol Sci* 356:1453–65. doi: 10.1098/rstb.2001.0945
- Mroczko-Wąsowicz A, Werning M (2012) Synesthesia, sensory-motor contingency, and semantic emulation: how swimming style-color synesthesia challenges the traditional view of synesthesia. *Front Psychol* 3:279. doi: 10.3389/fpsyg.2012.00279
- Mullally SL, Hassabis D, Maguire EA (2012) Scene construction in amnesia: an fMRI study. *J Neurosci* 32:5646–53. doi: 10.1523/JNEUROSCI.5522-11.2012
- Mullally SL, Vargha-Khadem F, Maguire EA (2014) Scene construction in developmental amnesia: An fMRI study. *Neuropsychologia* 52:1–10. doi: 10.1016/j.neuropsychologia.2013.11.001

- Müller GE, Pilzecker A (1900) Experimentelle Beiträge zur Lehre vom Gedächtnis. *Z Psychol Ergänzungsband* 1:1–300.
- Nádasdy Z, Hirase H, Czurkó A, et al (1999) Replay and Time Compression of Recurring Spike Sequences in the Hippocampus. *J Neurosci* 19:9497–9507.
- Nadel L, Moscovitch M (1998) Hippocampal contributions to cortical plasticity. *Neuropharmacology* 37:431–9.
- Nadel L, Moscovitch M (1997) Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr Opin Neurobiol* 7:217–27. doi: 10.1016/S0959-4388(97)80010-4
- Nadel L, Samsonovich A, Ryan L, Moscovitch M (2000) Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* 10:352–68. doi: 10.1002/1098-1063(2000)10:4<352::AID-HIPO2>3.0.CO;2-D
- Naqshbandi M, Roberts WA (2006) Anticipation of future events in squirrel monkeys (*Saimiri sciureus*) and rats (*Rattus norvegicus*): tests of the Bischof-Kohler hypothesis. *J Comp Psychol* 120:345–57. doi: 10.1037/0735-7036.120.4.34
- O’Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* 34:171–175.
- O’Keefe J, Recce ML (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3:317–30. doi: 10.1002/hipo.450030307
- Owen AM, Sahakian BJ, Semple J, et al (1995) Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia* 33:1–24.
- Pagin P, Westerståhl D (2010) Compositionality. *Semant. An Int. Handb. Nat. Lang. Mean.*
- Parsons T (1990) *Events in the Semantics of English*. MIT Press
- Patihis L, Ho LY, Tinggen IW, et al (2014) Are the “memory wars” over? A scientist-practitioner gap in beliefs about repressed memory. *Psychol Sci* 25:519–30. doi: 10.1177/0956797613510718
- Pavlidis C, Winson J (1989) Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *J Neurosci* 9:2907–18.
- Paxton R, Hampton RR (2009) Tests of planning and the Bischof-Köhler hypothesis in rhesus monkeys (*Macaca mulatta*). *Behav Processes* 80:238–46. doi: 10.1016/j.beproc.2008.12.016
- Payne DG (1987) Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychol Bull* 101:5–27. doi: 10.1037/0033-2909.101.1.5

- Paz R, Gelbard-Sagiv H, Mukamel R, et al (2010) A neural substrate in the human hippocampus for linking successive events. *Proc Natl Acad Sci U S A* 107:6046–51. doi: 10.1073/pnas.0910834107
- Pianesi F, Varzi AC (2000) Events and Event Talk. In: Higginbotham J, Pianesi F, Varzi AC (eds) *Speak. Events*. Oxford University Press, Oxford, pp 3–47
- Pyka M, Cheng S (2014) Pattern Association and Consolidation Emerges from Connectivity Properties between Cortex and Hippocampus. *PLoS One* 9:e85016. doi: 10.1371/journal.pone.0085016
- Raby CR, Alexis DM, Dickinson A, Clayton NS (2007) Planning for the future by western scrub-jays. *Nature* 445:919–921.
- Raffmann D (1995) On the Persistence of Phenomenology. In: Metzinger T (ed) *Conscious Exp. Schöningh/Imprint Academic, Paderborn*, pp 293–308
- Ranganath C, Blumenfeld RS (2005) Doubts about double dissociations between short- and long-term memory. *Trends Cogn Sci* 9:374–80. doi: 10.1016/j.tics.2005.06.009
- Ranganath C, Cohen MX, Brozinsky CJ (2005) Working memory maintenance contributes to long-term memory formation: neural and behavioral evidence. *J Cogn Neurosci* 17:994–1010. doi: 10.1162/0898929054475118
- Ranganath C, D’Esposito M (2001) Medial Temporal Lobe Activity Associated with Active Maintenance of Novel Information. *Neuron* 31:865–873. doi: 10.1016/S0896-6273(01)00411-1
- Ribot T (1881) *Les maladies de la mémoire*. Germer Baillare, Paris
- Ruchkin DS, Grafman J, Cameron K, Berndt RS (2003) Working memory retention systems: a state of activated long-term memory. *Behav Brain Sci* 26:709–28; discussion 728–77.
- Ryan L, Nadel L, Keil K, et al (2001) Hippocampal complex and retrieval of recent and very remote autobiographical memories: evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus* 11:707–14. doi: 10.1002/hipo.1086
- Ryle G (1949) *The Concept of Mind*. Hutchinson & Company, London
- Saksida LM, Bussey TJ, Buckmaster CA, Murray EA (2006) No effect of hippocampal lesions on perirhinal cortex-dependent feature-ambiguous visual discriminations. *Hippocampus* 16:421–30. doi: 10.1002/hipo.20170
- Schacter DL (2012) Constructive memory: past and future. *Dialogues Clin Neurosci* 14:7–18.
- Schacter DL (2002) *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Houghton Mifflin
- Schacter DL, Dodson CS (2001) Misattribution, false recognition and the sins of memory. *Philos Trans R Soc Lond B Biol Sci* 356:1385–93. doi: 10.1098/rstb.2001.0938

- Scoboria A, Mazzoni G, Kirsch I, Milling LS (2002) Immediate and persisting effects of misleading questions and hypnosis on memory reports. *J Exp Psychol Appl* 8:26–32. doi: 10.1037//1076-898X.8.1.26
- Scoville W, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20:11–21.
- Shanton K, Goldman A (2010) Simulation theory. *Wiley Interdiscip Rev Cogn Sci* 1:527–538. doi: 10.1002/wcs.33
- Sharon T, Moscovitch M, Gilboa A (2011) Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proc Natl Acad Sci U S A* 108:1146–51. doi: 10.1073/pnas.1005238108
- Shrager Y, Gold JJ, Hopkins RO, Squire LR (2006) Intact visual perception in memory-impaired patients with medial temporal lobe lesions. *J Neurosci* 26:2235–40. doi: 10.1523/JNEUROSCI.4792-05.2006
- Skaggs W, McNaughton B, Wilson M, Barnes C (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6:149–172. doi: 10.1002/(SICI)1098-1063(1996)6:2%3C149::AID-HIPO6%3E3.0.CO;2-K
- Skaggs WE, McNaughton BL (1996) Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science* 271:1870–3. doi: 10.1126/science.271.5257.1870
- Skaggs WE, McNaughton BL, Permenter M, et al (2007) EEG sharp waves and sparse ensemble unit activity in the macaque hippocampus. *J Neurophysiol* 98:898–910. doi: 10.1152/jn.00401.2007
- Squire LR, Alvarez P (1995) Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr Opin Neurobiol* 5:169–77.
- Squire LR, Zola-Morgan S (1988) Memory: brain systems and behavior. *Trends Neurosci* 11:170–5.
- Steinvorth S, Levine B, Corkin S (2005) Medial temporal lobe structures are needed to re-experience remote autobiographical memories: evidence from H.M. and W.R. *Neuropsychologia* 43:479–96. doi: 10.1016/j.neuropsychologia.2005.01.001
- Suddendorf T (2013) Mental time travel: continuities and discontinuities. *Trends Cogn Sci* 17:151–2. doi: 10.1016/j.tics.2013.01.011
- Suddendorf T, Corballis MC (1997) Mental time travel and the evolution of the human mind. *Genet Soc Gen Psychol Monogr* 123:133–67.
- 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; discussion 313–51. doi: 10.1017/S0140525X07001975

- Suddendorf T, Corballis MC (2008) New evidence for animal foresight? *Anim Behav* 75:e1–e3. doi: 10.1016/j.anbehav.2008.01.006
- Suzuki WA (2009) Perception and the Medial Temporal Lobe: Evaluating the Current Evidence. *Neuron* 61:657–666. doi: 10.1016/j.neuron.2009.02.008
- Toribio J (2007) Nonconceptual Content. *Philos Compass* 2:445–460. doi: 10.1111/j.1747-9991.2007.00075.x
- Toth JP, Hunt RR (1999) Not one versus many, but zero versus any: structure and function in the context of the multiple memory systems debate. In: Foster JK, Jelicic M (eds) *Mem. Syst. Process. or Funct. Debates Psychol.* Oxford University Press, New York, NY, USA, pp 232–272
- Trustwell R (2011) *Events, Phrases, and Questions.* Oxford University Press, Oxford
- Tulving E (1983) *Elements of Episodic Memory.* Clarendon Press
- Tulving E (1972) Episodic and Semantic Memory. In: Tulving E, Donaldson W (eds) *Organ. Mem.* Academic Press, Inc., New York, NY, USA, pp 381–402
- Tulving E (1985) Memory and Consciousness. *Can J Psychol* 26:1–26.
- Tulving E (1993) What Is Episodic Memory? *Curr Dir Psychol Sci* 2:67–70. doi: 10.1111/1467-8721.ep10770899
- Tulving E (1995) Organization of memory: Quo vadis. In: Gazzaniga M (ed) *Cogn. Neurosci.* MIT Press, Cambridge, MA, pp 839–847
- Tulving E, Kapur S, Craik FI, et al (1994) Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc Natl Acad Sci U S A* 91:2016–20.
- Wallenstein G V, Eichenbaum H, Hasselmo ME (1998) The hippocampus as an associator of discontinuous events. *Trends Neurosci* 21:317–23.
- Weiler JA, Suchan B, Daum I (2010) Foreseeing the future: occurrence probability of imagined future events modulates hippocampal activation. *Hippocampus* 20:685–90. doi: 10.1002/hipo.20695
- Weiss C, Bouwmeester H, Power JM, Disterhoft JF (1999) Hippocampal lesions prevent trace eyeblink conditioning in the freely moving rat. *Behav Brain Res* 99:123–32.
- Werning M (2012) Non-symbolic compositional representation and its neuronal foundation: towards an emulative semantics. In: Werning M, Hinzen W, Machery E (eds) *Oxford Handb. Compos.* Oxford University Press, Oxford, pp 633–654
- Werning M (2005) Right and wrong reasons for compositionality. In: Werning M, Machery E, Schurz G (eds) *Compos. Mean. Content.* Ontos Verlag, Frankfurt, pp 285–309

- Werning M (2010) Descartes discarded? Introspective self-awareness and the problems of transparency and compositionality. *Conscious Cogn* 19:751–761.
- Werning M (2003) Ventral versus dorsal pathway: The source of the semantic object/event and the syntactic noun/verb distinction? *Behav Brain Sci*. doi: 10.1017/S0140525X03400071
- Werning M, Hinzen W, Machery E (2012) *The Oxford Handbook of Compositionality*. Oxford University Press, Oxford
- Wheeler MA, Stuss DT, Tulving E (1995) Frontal lobe damage produces episodic memory impairment. *J Int Neuropsychol Soc* 1:525–36.
- Wilson MA, McNaughton BL (1994) Reactivation of hippocampal ensemble memories during sleep. *Science* 265:676–9. doi: 10.1126/science.8036517
- Wood ER, Dudchenko PA, Eichenbaum H (1999) The global record of memory in hippocampal neuronal activity. *Nature* 397:613–616. doi: 10.1038/17605
- Zentall TR, Clement TS, Bhatt RS, Allen J (2001) Episodic-like memory in pigeons. *Psychon Bull Rev* 8:685–90.
- Zola-Morgan S, Squire LR, Amaral DG, Suzuki W a (1989) Lesions of perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. *J Neurosci* 9:4355–70.