Title

Opening a conceptual space for metamemory experience

Authors

Neisser, Joseph

Abreu, George

Drane, Daniel L

et al.

Publication Date

2023-04-01

DOI

10.1016/j.newideapsych.2022.100995

Peer reviewed

HHS Public Access

Author manuscript

New Ideas Psychol. Author manuscript; available in PMC 2024 April 01.

Published in final edited form as:

New Ideas Psychol. 2023 April ; 69: . doi:10.1016/j.newideapsych.2022.100995.

Opening a conceptual space for metamemory experience

Joseph Neisser[a], **George Abreu**[a], **Daniel L. Drane**[b,d], **Nigel P. Pedersen**[c,e], **Thomas D. Parsons**[f], **Anne M. Cleary**[g]

[a]Department of Philosophy, Grinnell College, 1120 Park St, Grinnell, IA, 50112, USA

[b]Department of Neurology, Emory University, 101 Woodruff Circle, Atlanta, GA, 30322, USA

[c]Department of Pediatrics, Emory University, 101 Woodruff Circle, Atlanta, GA, 30322, USA

[d]Department of Neurology, University of Washington School of Medicine, Seattle, WA, 98105, USA

[e]Department of Biomedical Engineering, The Georgia Institute of Technology, Atlanta, GA, 30332, USA

[f]Simulation Science & Immersive Technology, Arizona State University, Tempe, AZ, 85281, USA

[g]Department of Psychology, Colorado State University, Fort Collins, CO, 80525, USA

Abstract

The experiences associated with remembering, including metamemory feelings about the act of remembering and attempts at remembering, are not often integrated into general accounts of memory. For example, David Rubin (2022) proposes a unified, three-dimensional conceptual space for mapping memory states, a map that does not systematically specify metamemory feelings. Drawing on Rubin's model, we define a distinct role for metamemory in relation to first-order memory content. We propose a fourth dimension for the model and support the proposal with conceptual, neurocognitive, and clinical lines of reasoning. We use the modified model to illustrate several cases, and show how it helps to conceptualize a new category of memory state: *autonoetic knowing*, exemplified by *déjà vu*. We also caution not to assume that memory experience is directly correlated with or caused by memory content, an assumption Tulving (1989) labeled the *doctrine of concordance*.

Keywords

Consciousness; Memory; Metamemory; Phenomenology; Déjà vu; Concordance

1. Introduction

Memory science must do more than construct and test models of information storage, retrieval, and reconstruction. A formidable task confronting any adequate account of

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

memory is to help explain memory experience - an essential aspect of human consciousness more generally. As Tulving (1989) warned and we discuss below, memory consciousness is not simply produced in "concordance" with information retrieval. Since then, significant progress has been made studying the so-called metamemory feelings associated with, but distinct from, first order functions (Dokic, 2014; Proust, 2013; Schwartz, 1999). But what is the relation between metamemory and base memory processing? How do the two processes interact to produce conscious recollection? How does taking metamemory more seriously impact our general understanding of memory?

Here, we offer a way to represent the distinct role of metamemory feelings in relation to first-order memory content. Whereas classical 20th century theories treat procedural, semantic, and especially episodic memory as encapsulated systems with distinct underlying mechanisms, more current approaches see them as drawing on diverse but partially overlapping neural substrates. Interestingly, metamemory feelings play a distinctive role in conscious recollection (Proust, 2013). Here we reconceptualize that role and synthesize the empirical, clinical, and theoretical evidence for it. Our approach promises new understandings of observed memory phenomena such as déjà vu (the feeling of reliving an experience when that seems impossible), jamais vu (the feeling of unfamiliarity with something that seems as if it should feel familiar), and other reported clinical cases. Our approach does not attempt a full explanation of the causal processes underlying memory experience. Instead, it provides a systematic conceptual map of the relations between various metamemory feelings and other aspects of information retrieval. In doing so, it builds on contemporary work by David Rubin (2022).

Efforts to renovate memory theory are well underway. For example, a special issue of Memory & Cognition was devoted to Rethinking the Distinction between Episodic and Semantic Memory (Mem Cognit 2022, Apr; 50(3). See also Michaelian, Debus, & Perrin, 2018; Klein, 2014; Madan, 2020). David Rubin is one important contributor to the contemporary effort to develop new models. Rubin (2022) proposes a unified dimensional model of memory states, wherein three axes — scene construction, self-reference, and explicitness — provide a single conceptual space within which, he argues, all memory states can be ordered. According to this framework, any memory state - procedural, semantic, episodic, emotional - can be characterized by these three kinds of information content. For example, episodic memory is depicted as a memory state consisting in explicit scene construction with self-reference, while semantic memory (knowledge of facts) is thought of as an explicit state without scene construction and without self-reference. Autobiographical narrative memory can be understood as explicit knowledge with self-reference but without scene construction. In this way, Rubin's model provides a useful vector map of the relations between memory states along the three dimensions. We introduce the dimensional model more fully in section 1, below.

However, and as Rubin (2022) partly acknowledges (pp.371), the dimensional model does not capture the phenomenology of memory because memory experiences are not direct reflections of the "first-order" memory processes such as storage and retrieval of procedural, semantic, or episodic information. Metamemory feelings, including the feeling of familiarity (FOF), the tip-of-the-tongue (TOT) state, déjà vu, and, we argue, the feeling of mental time

travel associated with episodic memory, cannot be mapped in the dimensional model as formulated. Because any adequate unified model must take into account the phenomenology of memory, we propose a friendly modification to Rubin's cube: the addition of a fourth axis of variation that captures the phenomenological dimension. We further speculate that the fourth dimension may represent a metacognitive marking or annotation function, alerting the animal that salient first-order information has been encountered before, and that the system is in retrieval mode (cf. Bartolomei et al., 2012). The paper has six short parts. In the first, we introduce Rubin's dimensional model and discuss the assumption of "concordance" between memory content and memory experience. In the second, third, and fourth parts, we introduce relevant ideas about déjà vu and metamemory from cognitive psychology, philosophy, clinical neuropsychology, and epileptology. The fifth and sixth parts introduce the proposed modification to Rubin's model and use it to illustrate a novel category of memory experience: autonoetic knowing.

2. The dimensional model of memory

David Rubin (2022) offers a new conceptual framework for memory states, within which all of the three traditional memory state types — procedural, semantic, and episodic — can be plotted. The state-space consists of three scalar dimensions: explicit/implicit, self-reference, and scene-specific contextual information. We call Rubin's approach the dimensional model, and informally refer to it as "Rubin's cube" (Fig. 2). The basic idea — with caveats, as we will see — is that any memory state can be understood as a function of these three variables. Traditionally, cognitive scientists have treated the three types of memory (procedural, semantic, and episodic) as if they are underwritten by three different neurocognitive systems, with three kinds of dedicated information. Diverse memories have not been understood as part of a single information system. Notice that the dimensional model is for memory states; it remains largely silent about etiology or functional role. That said, Rubin defends the choice of each dimension on both behavioral and neural grounds (pp.467). Thus, one strength of Rubin's model is that it is empirically informed but at the same time relatively theory neutral. Given the current state of flux in memory science, this is judicious. Rubin also holds that this relative neutrality helps create conceptual space for previously "homeless" memory phenomena that have resisted categorization according to the three primary memory types (episodic, semantic, & procedural), including déjà vu (pp.464). For instance, Rubin suggests that the dimensional model provides new conceptual space for phenomena such as: explicit scene memory without self-reference (for fiction and other people's memory), implicit scene memory with self-reference (for déjà vu), explicit non-scene memory with and without self-reference (for various neuropsychological cases), and lastly implicit, non-scene memories without self-reference (for aspects of personality disorder and phobias) (pp.473).1

Though it may possibly be true that any memory state of whatever type can be plotted onto the space defined by the three parameters of Rubin's cube, the plot does not capture some salient differences between memory states. That is, some states will occupy a virtually identical point in the dimensional space and yet they will be very different memory states. In particular, variations in memory experiences that are independent of the three kinds of memory content will be obscured by the model. Rubin acknowledges this point when he

says that episodic memory also includes “independent processes” that are not specified by the three dimensions, including especially an autonoetic phenomenology or feeling of mental time travel (pp.371). Rubin elaborates: “These added properties also remain part of the theory of episodic memory. The use of dimensions with added restrictions for other properties of episodic memory, instead of a single combined concept of episodic memory, should help clarify both the behavioral and neural basis of Tulving’s concept of episodic memory and how it fits into the theoretical organization of memory in general” (pp.372). We argue that these differences can be rendered intelligible, however, with the addition of a fourth dimension in the parametric space, the phenomenological dimension of memory feelings. Memory feelings can be cognitively modeled as metamemory annotations of the three-dimensional memory vector in Rubin’s cube, functioning to signal to the subject that the system is in retrieval mode. So the present paper assumes that the three axes of the dimensional model are necessary for any unified model of memory states, but argues that they are not sufficient to capture the phenomenology of recollection and other facets of memory, like sensations of familiarity or déjà vu.2 Thus, it seems a natural fit — a friendly amendment, so to speak — to suggest that the “independent” processes not specified by the dimensional model are those of metamemory: the noetic feelings of knowing and familiarity (FOK and FOF), as well as the autonoetic feeling of pastness (FOP) first identified by Russell, and 2013 (1921) and still thought to be crucial to the phenomenology of autonoesis (e.g. Dokic, 2014; Perrin, Michaelian, & Sant’Anna, 2020).

2.1. The doctrine of concordance and the dimensional model

Endel Tulving (1989) identified and criticized a common assumption among cognitive psychologists, which he dubbed the “doctrine of concordance.”

> “In order to deal explicitly with this tacit, unnamed assumption, it is first necessary to name it. I will refer to it as the doctrine of concordance of cognition, behaviour, and experience, or simply concordance. It holds that there exists a close and general, even if not perfect, agreement between what people know, how they behave, and what they experience. Thus, conscious awareness is required for, and therefore accompanies, the acquisition of knowledge, or its retrieval from the memory store; retrieved knowledge guides behaviour, and when this happens, people are aware of the relation between the knowledge and the behaviour; future behaviour is planned and ongoing behaviour is executed under the watchful eye of consciousness.” (Tulving, 1989, pp.8).

Two points are worth noting straight away: First, Tulving emphasizes a three-way concordance between knowledge, experience, and behavior. In this paper our focus is on the first two — the concordance (or lack thereof) between first-order memory content and memory experience. This is not an objection to Tulving, but simply because behavioral differences, in the form of verbal reports and observed neural activity, are cited as the evidence for lack of concordance between the first two. Second, Tulving’s critique was framed as part of a general push to rehabilitate consciousness as an object of study for psychology, understood as “the science of mental life” (Tulving, 1989, pp.5). He argued that because cognitive psychologists understand their remit as the study of information processes — the inner happenings that mediate and make possible thought and behavior

— they have tended to lose sight of the central and essential component of mental life, namely, experience itself. Tulving held that most cognitive psychologists, especially memory researchers, were simply unaware that they neglected to study memory experience. The field had a collective blind spot, generated by the tacit assumption of concordance. The language of memory theory consists of terms like search, scan, match, and recognize. This language, he suggested, invites an understanding in terms of conscious mental actions, making it easy to assume that they stand in concordance with experience (pp.7). Tulving reviewed several counterexamples to the concordance principle and concluded that there is no general concordance between behavior, knowledge, and experience when it comes to memory.

Schwartz (1999) followed up on Tulving's general diagnosis with a detailed review showing how concordance breaks down for the case of one particular metacognitive feeling, the tip-of-the-tongue experience or TOT. TOT feelings are not necessarily always caused by the same processes as word retrieval, but rather they can at least sometimes be caused by an inferential process derived from non-target information (pp.379. For an updated and somewhat modified version of this approach, see Schwartz, B. L., & Metcalfe, J., 2011). Over the past 20+ years metamemory has become a distinct subfield within memory psychology, in part as a result of the widespread recognition that memory feelings play a distinct set of functional roles in memory and must be explained by a distinct range of neurocognitive mechanisms.

Return, now, to Rubin's cube. The model tacitly assumes the concordance principle, i.e. that memory experience simply accompanies or is produced by the information content of the memory. In particular, the names for the three axes cry out for an interpretation in terms of consciousness: "scene construction" invites one to think that there is a scene presented to the mind's eye, as in an observer memory; "self-reference" invites the reader to suppose that the subject explicitly considers the content in relation to themselves, as in a first-person narrative; and "explicit" information is standardly used as a cognitive cypher for information that one is conscious of knowing. Thus construed, the three axes of the model suggest that the experience of mental time travel is given as a function of — in concordance with — these dimensions of knowledge or cognitive content.

This does not mean that Rubin's contribution is fatally flawed, however. On the contrary, it marks a significant step toward rethinking the episodic/semantic distinction and, as a result, toward understanding some of the remaining quirks of memory, such as déjà vu. We return to take up this thread below, after first introducing some necessary concepts from memory theory and clinical neuropsychology.

3. Cognitive and philosophical perspectives on memory experience

In this section we provide an indication of some of the recent work on metamemory feelings that tell against the Doctrine of Concordance and that should eventually push us beyond the three-dimensional state space provided in Rubin's model. Eventually it will push us, in sections 5 Neurological evidence of discrete neural circuits for metamemory "annotations", 6 The dimensional model and metamemory feelings, to propose a fourth dimension for the model, mapping the activity of metamemory.

Psychologists have identified and manipulated several feelings associated with conscious retrieval of semantic knowledge, including the feeling of knowing (FOK), the tip-of-the-tongue (TOT) state, the feeling of familiarity (FOF), and confidence (e.g., Dunlosky & Tauber, 2016). These noetic feelings are widely thought to be functional, guiding cognitive actions such as continued memory search (Huebert et al., 2022; Schwartz, 2002; Schwartz & Cleary, 2016), predictions about future performance, or decisions about whether to complete a task. But noetic feelings are subpersonal or procedural forms of metacognition that do not depend on a rich self-concept or self-awareness (Proust, 2013). In contrast, episodic memory is "autonoetic" because it involves no mere feeling of familiarity or knowing but instead a distinct feeling of pastness (FOP), or what we referred to just above as the feeling of mental time travel. The FOP is a feeling that the episode in view took place in my personal past, that it has an "earlier" location on my subjective timeline. The FOP is thought to be richer than the FOF because its content includes awareness of the temporally ordered experiences of a stable and continuing self. In this way, the concept of the FOP is also reminiscent of the feeling of "me-ness" that was suggested by Claparede to be involved in some, but not all, sensations of familiarity (Kihlstrom, 1995).

3.1. Autonoetic versus noetic awareness

Whereas episodic memory deals with specific episodes or events in time, semantic memory involves knowledge that is more generic and abstract. The distinction between semantic and episodic memory was initially a heuristic one, aimed at distinguishing these different aspects of memory based on the types of information represented and retrieved (Tulving, 1972). Tulving's heuristic distinction then led to research on whether the two facets of memory are fully independent (e.g., Devitt, Addis, & Schacter, 2017; Park, Miller, Nili, Ranganath, & Boorman, 2020). Later, the distinction evolved to include types of consciousness and metamemory feeling (Gardiner, 2001). Episodic memory came to be associated with autonoetic consciousness and semantic memory came to be associated with noetic consciousness. It was then that episodic memory came to be understood as mental time travel.

Episodic memory displays autonoetic phenomenology that differs from the "just knowing" that is found in semantic memory (Tulving, 1983). Gardiner (2001, p. 1351) held that this type of memory involves a "sense of self-recollection in the mental re-enactment of previous events at which one was present." There is a first-person aspect to autonoetic consciousness —a sense of oneself at a particular point in space and time, as if engaging in mental time travel through mental re-enactment of the past experience. Note that mental time travel might not only take one into the past, but also to the future insofar as one can imagine one's place in space and time in possible future or counterfactual scenarios (e.g., De Brigard, 2014; Schacter, Addis, & Buckner, 2007; Szpunar, Spreng, & Schacter, 2014).

Autonoetic phenomenology is meant as a contrast with noetic phenomenology, in which information from memory is brought to consciousness without an accompanying sense of re-experiencing the episode in which it was acquired (Gardiner, 2001). Examples include retrieving facts from memory or bits of knowledge that have been abstracted away from the place and time at which they were previously encountered, or simply knowing something

without remembering specifically when or where that knowledge was learned. Hence, noetic phenomenology is associated with semantic memory, and sometimes called “just knowing.” Just knowing something does not involve a subjective experience of mental time travel because it is a simpler act of bringing information back to consciousness.

3.2. A metacognitive approach to autonoetic experience

“We may say, then, that images are regarded by us as more or less accurate copies of past occurrences because they come to us with two sorts of feelings: (1) Those that may be called feelings of familiarity; (2) those that may be collected together as feelings giving a sense of pastness. The first leads us to trust our memories, the second to assign places to them in the time-order.” - Bertrand Russell, and 2013 (1921), Lecture IX, “Memory.”

A promising metacognitive approach to the FOP has been proposed by Jérôme Dokic (2014; see also Perrin et al., 2020; Isingrini et al., 2016, Proust, 2013).3 Dokic identifies a feeling of first-handedness, and argues that this is what distinguishes autonoetic from noetic forms of memory: “An episodic memory is not merely first-hand, but it subjectively feels first-hand. In other words, it feels to originate directly from one’s past experience, excluding the essential involvement of either reasoning or testimony” (Dokic, 2014, pp.416, emphasis original). First-handedness, he argues, is a distinct kind of feeling of knowing (FOK), an episodic FOK or EFOK. The EFOK is what would have been missing in the case of RB (discussed below), who lost the sense of “personal own” despite having an otherwise intact episodic recall.

Dokic’s proposed EFOK is functionally different from the noetic or semantic FOK (dubbed the SFOK) because it is sensitive to “scene” information (to use Rubin’s terminology): “The implicit monitoring mechanisms underlying EFOKs are sensitive to a variety of cues over and above familiarity, such as the presence of some perceptual, spatio-temporal and/or ‘gist’ information, or whether the subject is spontaneously searching for extra detail” (Dokic, 2014, pp.421). The subject then uses the EFOK as a signal to distinguish whether the constructed episode is from memory or whether it is a hypothetical or counterfactual imagining. That is, it signals when the system is in retrieval mode. Apparently, though, the mechanism is highly imperfect - it can often trigger in false memory contexts, and it can even trigger in absence of any episodic content (see below).

The theory of the EFOK is more cognitively minimal than the classical feeling of pastness because it does not require mastery of the high-level concepts of self or fixed temporal succession – hence it is a variety of “just knowing.” But it is a knowing that is triggered by pattern recognition of events and layouts. It is an annotation function rooted in the cognitive map. This idea can have great utility for the explanation of déjà vu, in which there is an episodic version of “just knowing,” a category of memory state that we think of as autonoetic knowing (see below).

The takeaway point at present is that the sense of “reliving” a scene, prominent in both episodic memory and déjà vu, and which seems to escape the dimensional model, may well be generated by procedural metacognition, in much the same way that semantic memory is often accompanied by noetic feelings of familiarity. In a similar vein, Madan (2020) has

suggested that the feeling of familiarity, usually associated with semantic memory, may also be at the root of episodic memory. If this line of reasoning is on track, then (some forms of) episodic memory may well be evolutionarily older than previously supposed, phylogenetically prior to sophisticated concepts of self or temporal order, and widely shared across the animal kingdom. Metamemory, then, is a fundamental process that must be part of any general account of memory, including the dimensional model.

4. Clinical perspectives on memory experience

Much of the received wisdom and a large body of research about the relation between episodic and semantic memory originally stems from the case HM, in which doctors undertook a nearly complete bilateral hippocampotomy. Today, more minimally invasive neurosurgical techniques have become available that are typically only performed on one hemisphere (Bourdillon et al., 2020; Liscak et al., 2010; Willie et al., 2014). Rather than a large, en bloc section of tissue being resected from the temporal lobe (Ojemann & Valiante, 2006), neurosurgeons can now ablate very focal regions of tissue while sparing much of the overlying cortex and neural pathways through the use of laser interstitial thermal therapy (LITT) and/or radiofrequency (RF) ablation (Bourdillon et al., 2020; Willie et al., 2014). These techniques are allowing for a “renaissance” of lesion analysis studies in the context of epilepsy or tumor surgery (Donos et al., 2018; Drane, 2018). Clinical experience is evolving accordingly, and now cognitive theory must also update. In this section we review the changing understanding of the case HM as well as another relevant clinical example, the case of RB.

4.1. Reassessment of HM

The case of HM led to significant changes in the practice of epilepsy surgery, and crystallized the seeming importance of the hippocampus as the central structure necessary for making novel episodic memories. The initial article by William Scoville (neurosurgeon) and Brenda Milner (neuropsychologist) was entitled, “Loss of recent memory after bilateral hippocampal lesion,” and this title reflects the presumed importance of the hippocampus (Scoville & Milner, 1957). The tragic outcome in the case of HM led to the introduction of new techniques, such as the intracarotid amobarbital Wada procedure, to determine which hemisphere was important for memory, operationalized as visual recognition memory and naming. Additionally, the critical role of the hippocampus in memory has been solidified by later evidence that anoxic injury, disproportionately affecting the metabolically intensive hippocampi, can lead to a similar amnestic state (Gadian et al., 2000; Hopkins & Haaland, 2004; Vargha-Khadem et al., 1997). Hippocampal atrophy, as measured through volumetric analysis in a variety of patient groups (e.g., Alzheimer’s, epilepsy, hypoxic injury), has also been correlated with poor memory function (Allen, Tranel, Bruss, & Damasio, 2006; Gorbach et al., 2020; Issacs et al., 2003; Vargha-Khadem et al., 1997).

Nevertheless, the original manuscript of HM actually noted that damage was much more extensive than just the bilateral hippocampi (Scoville & Milner, 1957). Once neuroimaging technologies were developed, it became possible to study the extent and location of surgical lesions created in the brain of HM more precisely (Corkin, Amaral, Gonzalez, Johnson,

&, Hyman, 1997), and, eventually, post-mortem histopathology (Annese et al., 2014). It has become clear that the resection involved almost total destruction of the bilateral entorhinal cortices, and included the piriform and parahippocampal cortices, the medial temporal pole, and a very small amount of anterolateral temporal lobe. The resection of the amygdala and hippocampal formation was less than initially reported, but did include most of the amygdaloid complex and approximately half of the rostro-caudal extent of the intraventricular segment of the hippocampal formation (including dentate gyrus, hippocampus, and subiculum). The resection was described in Scoville's report as extending about 8 cm from the temporal tip but was actually more in the range of 5 cm. During the post-mortem brain cutting, it was also discovered that HM had suffered a small lesion involving the left lateral orbitofrontal gyrus and the underlying white matter. Of note, Scoville and Milner described other bilateral hippocampal surgical resection cases, and it appears that several of the others were not rendered amnestic by these surgeries, although some experienced mild to moderate memory problems, and the reported lesions extended a lesser distance from the temporal tip (Scoville & Milner, 1957).

This reevaluation suggests that HM ′s amnestic pathology was not a result merely of hippocampal damage, but there were likely contributions from surrounding regions, consistent with non-human primate literature (Zola-Morgan, Squire, & Ramus, 1994). Additionally, another recent and relevant case report claims to have successfully performed a radiofrequency ablation of the bilateral hippocampi in a single individual who also did not become amnestic and did not significantly decline on the Chinese version of the Wechsler Memory Scale (which was fairly poor at baseline: Luo et al., 2013) implying that more may be at play in memory networks than previously assumed. All of these reports together suggest that the paradigmatic schema of the hippocampus as the chief component of multiple types of memory, as well as the clean functional distinction between episodic and semantic memory, need to be updated.

4.2. A possible contrasting case: RB

For the present discussion it is also worth noting the case of RB, described by Klein and Nichols (2012). It should be stressed that the total amount of information and clinical observation about RB is limited, and that the facts of the case are not fully established. Accordingly, we do not make this case integral to our larger argument. Instead, we introduce it along with another related case (and related data about jamais vu) as part of our general discussion of memory feelings.4 But the case is there in the literature, and serves as a thought-provoking contrast to that of HM. While HM could recall his remote past, and semantic facts about the world, he could not form immediate new conscious memories. In contrast, RB eventually recovered scene memories of the accident and the time that followed, but still lacked any feeling of "personal ownership," that is, a sense that the episodes had happened to him. This patient, then, reports a pattern of experience that suggests something like inverse déjà vu: he did not feel as if the episode was from his past experience, even though he could call the scene to consciousness in a detailed way. In sections 5 Neurological evidence of discrete neural circuits for metamemory "annotations", 6 The dimensional model and metamemory feelings below, we try to reconceptualize this

case as an example of “dis-autonoesis,” and compare it with another memory quirk known as jamais vu, or “never seen” (cf. Moulin, Bell, Turunen, Baharin, & O’Connor, 2021).

RB is described as a 43-year-old, Caucasian, male who incurred a serious head injury when hit by a car while riding a bicycle (Klein & Nichols, 2012). RB showed cognitive effects including attentional deficits, mild aphasia, retrograde and anterograde amnesia for the events in close proximity to the injury. Several months later, RB began to recover memories for some of these pre- and post-injury events, and his neurocognitive deficits generally improved as well. As RB recovered, he gained recall for the temporal and spatial context of the events before and after the injury, and had intact self-referential knowledge. That is, he could conjure the scene and he knew, intellectually, what event it represented. This case, then, is suggestive. It seems possible (though by no means certain) that what RB lacked was the infamous feeling of mental time travel. RB knew that the episode had happened to him, but he no longer felt that it had. Klein &and Nichols expressed this by saying that RB had intact episodic recollection in the absence of a “personal own” for these memories.

Klein and Nichols (2012) also noted a similar case study published by Stuss and Guzman (1988). In that case a patient semantically relearned his personal history after incurring severe retrograde episodic amnesia for most of his life history. The patient reported that the memories seemed foreign or impersonal. Ultimately, they were able to learn the details but without recall for temporal and spatial context, not as scenes that were conjured before the mind’s eye. This might contrast with RB, in which scene memory was apparently recovered (including spatiotemporal context and imagery), only without the sense of “personal own.” Below (sections 5 Neurological evidence of discrete neural circuits for metamemory “annotations”, 6 The dimensional model and metamemory feelings) we argue that these cases, along with jamais vu, suggest that it is possible to dissociate scene memory from the principal markers of autonoetic experience, the “feeling of pastness” (FOP) or the episodic feeling of knowing (EFOK) (Dokic, 2014).

William James (1890) referred to a feeling of “warmth and intimacy” that goes with explicit recollective content, and which may be necessary if we are to weave scene memory into our personal sense of self. Déjà vu, then, is the presence of this episodic feeling — in which the individual feels personally connected to an episode and perhaps even experiences a sense of reliving the episode — in the absence of calling to mind any specific memory content for the feeling. The case such as RB suggests that the opposite phenomenon is also possible: There can be a calling to mind of scene content from memory without the episodic feeling. If so, the subcomponents of recollective experience can be dissociated.

Again, we stress that our larger approach does not stand or fall with the case of RB. But it illustrates how clinical experience since HM has gradually moved practitioners beyond the assumptions about the organization of memory systems that were initially grounded in that case. The cognitive and philosophical theory descending from HM has not kept in touch with accumulating evidence, both clinical and neurobiological. We turn to the latter in the next section.

5. Neurological evidence of discrete neural circuits for metamemory “annotations”

Neurologists have long distinguished epileptic déjà vu from common déjà vu by its intense character and other often stereotyped subjective symptoms, and several rigorous studies have appeared in the last decade (Adachi et al., 2010; Illman, Butler, Souchay, & Moulin, 2012; Warren-Gash & Zeman, 2014). While “déjà vu” may be an umbrella category covering several varieties of memory experience, epileptic déjà vu has been crucial to understanding the neural correlates of déjà vu more generally.5 We further argue that epileptic déjà vu also provides a setting in which associated feelings of familiarity and other quirks of consciousness can be studied. There is perhaps no greater contributor to our present understanding of these phenomena, and to future research in metamemory, than through the unique and direct access afforded to the human brain while planning surgery for patients with medication-refractory epilepsy. This section reviews some of the neurological evidence pointing to stratification of memory and metamemory function, especially as they give rise to memory experiences.

5.1. Rhinal cortices as neural substrates for metamemory feelings

The ability to elicit déjà vu via direct stimulation has important implications. Firstly, electrically evoked déjà vu, in its purest form, is devoid of mnemonic content (cf. Gloor, 1997, pp.701–702). Secondly, the method of stimulation mapping is fundamentally about localization, such that a clinical phenomenon may be attributed to specific parts of the nervous system. But traditional localizationalist accounts - attributing a function to a specific cortical regions - have largely been replaced by an understanding that everyday cognitive functions better correspond to large scale brain circuits. Accordingly, stimulation mapping is today understood to typically affect a large circuit rather than a specific locale. A classical example might be the speech arrest and impairment of comprehension that occurs during stimulation mapping of canonical parts of the language network, instead of the development of more complex aphasic syndromes that might occur with a lesion to one of these regions. This circuit approach to localization corroborates the findings, discussed above, of peripheral hippocampal regions contributing to the verbal and other aspects of memory. Thus, the ability to evoke pure déjà vu through stimulation of specific neural circuits suggests that there may be a particular anatomical substrate for the déjà vu feeling that is discrete from first-order content. By extension, it seems likely that the feeling of familiarity is also localizable and discrete from the cognitive content that it often accompanies. In other words, the brain annotates some cognitive contents with a metacognitive feeling of familiarity, to indicate when the system is in a ‘retrieval’ mode’ (Bartolomei et al., 2012). If there is a specific physiological process or anatomical circuit recruited for this function, then it may become active independently of the processes that access or construct the memory content. (In fact, it is even possible that the annotation function helps to signal to other cognitive networks the need to switch into retrieval mode to conduct a search of memory through inward directed attention). When this happens, the subject may mistakenly infer that an ongoing experience is being relived or recalled, i.e. they may experience déjà vu. While this possibility is compelling, particularly in light of likely discrete circuits for retrieval and the formation of episodic (or autonoetic) memory

(Roy et al., 2017), how it occurs is unknown. It seems likely that we will make progress by examining spontaneous, or externally-induced or perceptual déjà vu, for example with videographic or visual stimuli, in patients undergoing invasive electrophysiology. In line with this inference, the occurrence of déjà vu has been noted as more likely when electrical stimulation of the rhinal cortex results in increased synchrony between this structure and amygdala and hippocampus (Bartolomei et al., 2012). Most crucially, Bartolomei et al. state that "stimulation of the RCs (perirhinal and entorhinal cortices) most reliably triggered déjà vu in epileptic patients (Bartolomei et al., 2004) . . . Such results appeared [to] concur with previous assumptions (Spatt, 2002), suggesting that the appearance of déjà vu is due to the transient alteration of a 'familiarity system.' '(Bartolomei et al., 2012, pp.6–7). Bartolomei et al. present the hypothesis — based on previous studies on déjà vu — that the data can be plausibly interpreted as the system being in 'retrieval' mode (ibid.). This suggests that associations between the rhinal cortex and the hippocampal region, alongside others in the widespread medial temporal lobe (MTL), provide a plausible neural substrate for the phenomenon of déjà vu and its "familiarity" annotation system. Ongoing contemporary research suggests that the "collision of encoding and retrieval" states of this entorhinal, hippocampal and rhinal network underlies déjà vu (Gillinder, Liegeois-Chauvel, & Chauvel, 2022). Of course, key questions remain about how these observations fit and can contribute to philosophical and cognitive psychological views of memory and the conscious experience of memory.

5.2. Implications of neurological observations

While the above remains speculative, it seems reasonable to make some comments about accounts of memory. One example would be classes of mathematical models of recognition memory (e.g., Clark & Gronlund, 1996) that stem from a signal detection approach (e.g., Atkinson & Juola, 1974) to the computation of the familiarity signal. This general approach to familiarity — pitched at the functional or cognitive level — assumes that familiarity is a continuum, with novel situations or stimuli having very low familiarity. This seems difficult to reconcile with neural accounts of déjà vu, especially in the light of the delimited onset and end of déjà vu experiences. Furthermore, there is evidence that the neural circuits involving novelty include discrete hypothalamo-hippocampal circuits that even segregate modalities such as context versus social novelty (Chen et al., 2020), with the latter including a population of supramammillary neurons, some of which release both GABA and glutamate, that target a specialized portion of the hippocampus (Pedersen et al., 2017). These studies highlight the neural intricacy of memory circuits in relation to particular context and functions. These mechanisms appear to be very different from the known neural substrates of déjà vu, and suggest a more ethologically inspired model of circuits for specific contexts, rather than a general computational processor that encodes novelty and familiarity as a continuous variable. In short, déjà vu, while still a neurological puzzle, presents a potential window from which to investigate the neural circuitry involved in metamemory experience.

A final remark drawn from the above is that key phenomena such as déjà vu and familiarity – formerly marginal in the study of memory – offer powerful new evidence that metamemory processes have been underemphasized and underappreciated in the literature on memory. New research directions would do well to place more attention on integrating

metamemory into classical ethological and cognitive approaches. It may also be noted that a turn to metamemory resembles the beginnings of a rise of affectivism in neuropsychological research, both focusing on the functional and explanatory importance of minimally cognitive emotional-affective feelings in accounts of mental life (Dukes, Abrams, Adolphs et al., 2021).

6. The dimensional model and metamemory feelings

With all of the above in place, we return to Rubin's cube and our promised four-dimensional version of the model. As Rubin recognized, the feeling of mental time travel is not adequately represented as a three-place relation between explicitness, self-reference, and scene information. This is because it is possible to have all three but still not have an autonoetic phenomenology. For example, I can remember the kitchen layout, recall that the cupboard is next to the sink, and know that my knowledge was acquired through personal experience, all without any sense of "reliving" an episode of being in my kitchen. Conversely, I can indeed have the autonoetic phenomenology even while clocking very low on all three of the dimensions in the model. That is what happens in the case of déjà vu. This section introduces our proposed modification to Rubin's cube: a fourth dimension corresponding to metamemory feelings.

Our suggestion, then, is that the addition of a fourth axis, representing metamemory, better fulfills the dimensional model's goal of a unified conceptual space that can map any memory state in a single vector.

Rubin's claim to have found a new "home" for déjà vu in the cubic space follows only on the assumption of the Doctrine of Concordance. That is, only if the autonoetic phenomenology of déjà vu is to be explained as a function of the three axes in the model. But déjà vu is an intense feeling that is as of an episodic memory, but is otherwise empty at the first-order level of memory content. Roughly the opposite holds for jamais vu and in particular RB, the latter reported a similar state to jamais vu. That is, RB's memory state fulfills all three axes for episodic memory — it is an explicit scene memory with self-reference — but lacks the feeling of mental time travel in accordance with the alienation characteristic of jamais vu proper.

These two cases, déjà vu and jamais vu, cannot be adequately represented with Rubin's three dimensions, since Jamais vu (or RB) will not be distinguishable from episodic memory, while déjà vu will not appear at all. Like the TOT state documented by Schwartz (1999), these two kinds of memory experience violate the assumption of concordance between what we know and what we experience.

6.1. Opening the conceptual space for memory feelings

To open the way to a new dimension for Rubin's cube, we need to show that the new "space" is orthogonal to the three existing dimensions, i.e that there can be variations along the 4th axis that are not simply correlated with variations along the first three. Traditionally, the assumption of concordance pushed theorists to treat the episodic/semantic distinction as equivalent to the autonoetic/noetic distinction. Basically, writers proceeded as though

the two distinctions pick out the same phenomena. Table 1, Table 2 illustrate the point. By distinguishing between these two distinctions, we show how metamemory experience can occur apart from first-order memory content. In particular, Table 2 shows that déjà vu can be understood as "autonoetic knowing," while jamais vu and cases like RB are seen involve a kind of "dis-autonoesis." The existence of the new categories shows the need for the new axis on the dimensional model (see Fig. 1), providing a new conceptual space for "homeless" memory states such as déjà vu and jamais vu. In distinguishing between distinctions in this way, we adopt a method long used by philosophers including Kripke (1972), Williams (1953), and Kant (2008 [1781]). Table 1 exemplifies the assumption of concordance, criticized by Tulving and Schwartz. Table 2 shows what it looks like to abandon that assumption.

Table 2 shows a non-empty set at the intersection of semantic memory and autonoetic phenomenology, the new category of autonoetic knowing, which we playfully describe as "just reliving." Autonoetic knowing takes place when you have the sensation of mental time travel, but no actual information about what, where, or when, you are traveling to. This captures the surprising point that the déjà vu experience has much in common with the experience of semantic memory – it is a kind of knowing that, i.e., it is the knowledge that this has happened before.

Meanwhile, the intersection of episodic memory with noetic feeling now contains a broad category labeled "dis-autonoesis." Here there are states with episodic content but non-paradigmatic metamemory experiences — either a distortion or a complete lack of feelings of mental time travel. In these cases, the subject recognizes that the information has been encountered before and that it should feel familiar in a certain way, but experiences it differently. Moving beyond the suggestive but obscure case of RB (discussed above in section 3.2), a better studied 'correlate' of déjà vu is jamais vu, or "never seen." Jamais vu describes the opposite kind of distortion of familiarity – that is, when a well-known and explicitly recognized object (e.g. a word, item, or scene) feels oddly unfamiliar or strange. Jamais vu is often less intense than déjà vu, which may be one reason why it is less often reported. But psychiatrists sometimes consider jamais vu under the rubric of depersonalization (e.g. Sno, 1994; Silbermann, 1963; Devereux, 1967). In more recent experimental research, Moulin et al. (2021) used a word satiation paradigm in which subjects were instructed to continuously write a selected English word for up to 1 min or until they felt 'peculiar.' A significant proportion of subjects developed a sense of derealization as the words 'lost meaning' from continuous repetition. The authors argued that this alteration of memory phenomenology from its first-order memory content – i.e., the alienation of the metamemory feeling - is the essence of jamais vu (Moulin et al., 2021). Here, we take our cue from these authors, and apply "jamais vu" to those memory experiences that share this 'unifying mismatch' – i.e., episodic recognition and/or recall but a missing or distorted metamemory feeling to go along with it. In Table 2, this is placed in the conceptual space for "dis-autonoesis" along with RB.

The caveat in interpreting Table 1, Table 2 is that our simple 2×2 analysis is not meant to capture all the variation that would be necessary to represent all the facts about phenomena like jamais vu and déjà vu. Rather, it illustrates ways in which variation in metamemory

feeling can disrupt the assumption of concordance, and it thereby opens the conceptual space for a fourth dimension of Rubin's cube.

We now turn to mapping our new categories onto the four-dimensional space of Rubin's hypercube. To represent the fourth dimension, we have added a color scale. A "hot" memory state accompanied by a metacognitive feeling is represented in red. A "cool" state, unaccompanied by autonoetic phenomenology, is represented in blue.

7. Mapping metamemory

In this section, we illustrate the 4-dimensional model for memory states with three test cases: paradigmatic episodic memory, déjà vu, and jamais vu.

We begin with a paradigmatic episodic memory (see Fig. 2). In the 4-dimensional model, the vector plot within Rubin's three first-order cognitive axes is exactly the same as in the original model: High on scene construction, self-reference, and explicitness. And a typical example of episodic memory will also hit moderately high on the fourth dimension of metacognitive feeling. Typically, episodic memory states are marked (or "annotated") with an episodic feeling of knowing, the EFOK (see section 2.2, above). The metacognitive feeling may function to alert the subject that the episode being entertained is a memory rather than a hypothetical or counterfactual state that is merely imagined (Dokic, 2014).

Fig. 3 represents a typical case of déjà vu. The 4-dimensional model can represent this case as follows. Along the three first-order cognitive axes, déjà vu plots high in self-reference and explicitness but low (though non-negligible) on the scene-construction axis.6 But the example case of déjà vu is positioned very high ("hot") in the metacognitive feeling axis, reflecting the episodic feeling of knowing or EFOK that is an essential part of the state. This representation simultaneously captures the characteristic combination of a salient phenomenology with a lack of mental scene construction that is typical of déjà vu.

Note that although Rubin (2022) suggests that déjà vu is a case of implicit scene memory, we suggest that it is an explicit state. Our suggestion stems from the subjective phenomenology of the feeling of reliving an episode during déjà vu, and the fact that the feeling itself is conscious. It tends to pull the experiencer's attention away from whatever they were in the midst of prior to the feeling, toward the feeling, and to consciously prompt a memory search; thus, although the person is able to reconstruct the prior scene(s) potentially responsible for the déjà vu state, they are explicitly aware of the feeling itself. This is illustrated by the fact that the experience of déjà vu is often encoded into memory in a way that is consciously retrievable later on, as evidenced by the fact that people can report on their previous déjà vu experiences in survey research (e.g., Cleary & Brown, 2022). These are positive reasons for plotting déjà vu high in explicitness (see also Aitken & O'Connor, 2020). Further, there is reason to think that déjà vu should not be plotted on the implicit end of the scale, because it does not fit with other cases of implicit memory in which an unconscious content guides behavior in an unacknowledged way. In the classic example, Claparede reported on people who can acquire new procedural and semantic knowledge - e.g. that some people carry pins in their hands - even without being able to form new

episodic memories of the occasion on which they acquired it (Kihlstrom, 1995). That knowledge, then, is "implicit." But déjà vu does not constitute any comparable knowledge that can guide behavior in a way that is unknown or unacknowledged by the subject. In fact, people can believe that their déjà vu carries knowledge of what will come next while being mistaken about this (Cleary & Claxton, 2018). Hence, because the experience is highly salient and reportable, but carries no implicit procedural or semantic content, we plot it high on the explicitness axis of the dimensional model.

Fig. 4 presents the memory state of jamais vu as depicted in Moulin et al. (2020) word satiation paradigm. The case is plotted high in self-reference to capture the 'peculiarity' in memory experience which was selected for in those who reported. This placement follows from the study's characterization of jamais vu as consciously happening to oneself, as opposed to the earlier accounts of jamais vu as a form of depersonalization (see Sno, 2000). On the axis of scene construction, this example of jamais vu is plotted near the zero-point primarily because this mental state is not spatiotemporally dependent.7 Finally, for the third primary axis, jamais vu's explicitness is above the midpoint to do justice to the chosen paradigm's 'peculiarity' condition which is what brings attention to the increasingly dissonant mental state first. The 'peculiarity' itself (or the memory-metamemory dissonance) is a result of the phenomenological valance being rather low, as jamais vu's character is precisely the experience of knowing a memory item but still having no 'grounding feeling' to go along with it. Fig. 5.

8. Conclusion

Adding a fourth dimension of variation to Rubin's cube allows a systematic representation of the similarities and differences among metamemory states including but not limited to déjà vu. The important role of metamemory feelings has not been fully appreciated in general accounts of memory (though it has become a thriving subfield of special investigation). Converging lines of reasoning from neurology, clinical neuropsychology, cognitive psychology, and philosophy all point to a distinctive contribution of metamemory in the construction and the experience of memory states. Here we attempt to locate metamemory feelings in relation to first-order memory content, using as our point of reference Rubin's three-dimensional conceptual space for memory states. We have shown that a newly expanded conceptual space that includes a metamemory dimension can help achieve Rubin's goal of a truly unified representational map of all possible memory states. Immediate implications include: (1) that there may be an "annotation" function for metamemory feelings, signaling that the system is in retrieval mode rather than in an entirely novel or counterfactual mode; (2) that there are categories of memory states, which we dub "autonoetic knowing" and "dis-autonoesis," respectively, in which there is either an experience of mental time travel without episodic memory content, or vice-versa. Our approach also suggests several lines of future empirical and clinical research designed to identify the potential neural circuits for familiarity and the episodic feeling of knowing, or EFOK. Such research might contribute to the understanding of large-scale scientific problems such as the contribution of memory to social-emotional processing, or the relation between memory and ongoing conscious experience.

Funding and Acknowledgements

Dr. Drane's efforts on this paper were supported in part by funding received from the National Institutes of Health/National Institute of Neurological Disorders and Stroke, NIH/NINDS, R01NS088748). He also receives funding from additional grants from the NIH (R01MH118514; R01NS110347), as well as Medtronic, Inc. (A1321808). Dr. Pedersen is supported in part by funding received from the NIH/NINDS (K08NS105929, R21NS122011) and CURE Epilepsy. [4] [5].

Data availability

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

References

1. Adachi N, Akanuma N, Ito M, Adachi T, Takekawa Y, Adachi Y, et al. (2010). Two forms of déjà vu experiences in patients with epilepsy. Epilepsy and Behavior, 18, 218–222. 10.1016/j.yebeh.2010.02.016 [PubMed: 20494621]
2. Aitken CB, & O'Connor AR (2020). Converging on an understanding of the déjà vu experience. In Schwartz, & Cleary (Eds.), Memory quirks: The study of odd phenomena in memory (pp. 288–305). Routledge.
3. Allen JS, Tranel D, Bruss J, & Damasio H (2006). Correlations between regional brain volumes and memory performance in anoxia. Journal of Clinical and Experimental Neuropsychology, 28(4), 457–476. 10.1080/13803390590949287 [PubMed: 16624778]
4. Annese J, Schenker-Ahmed NM, Bartsch H, Maechler P, Sheh C, Thomas N, et al. (2014). Postmortem examination of patient H.M.'s brain based on histological sectioning and digital 3D reconstruction. Nature Communications, 5, 3122. 10.1038/ncomms4122
5. Arango-Munoz. (2014). The nature of epistemic feelings. Philosophical Psychology, 27(2), 193–211.
6. Arango-Munoz, & Bermudez. (2018). Remembering as a mental action. In Michaelian D, & Perrin (Eds.), New directions in the philosophy of memory. Routledge Press.
7. Atkinson RC, & Juola JF (1974). Search and decision processes in recognition memory. In Krantz DH, Atkinson RC, Luce RD, & Suppes P (Eds.), Contemporary developments in mathematical psychology: I. Learning, memory and thinking. W. H. Freeman.
8. Bartolomei F, Barbeau EJ, Nguyen T, McGonigal A, R ´egis, J., Chauvel, P., et al. (2012). Rhinal-hippocampal interactions during déjà vu. Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology, 123(3), 489–495. 10.1016/j.clinph.2011.08.012
9. Bourdillon P, Rheims S, Catenoix H, Montavont A, Ostrowsky-Coste K, Isnard J, et al. (2020). Surgical techniques: Stereoelectroencephalography-guided radiofrequency-thermocoagulation (SEEG-guided RF-TC). Seizure, 77, 64–68. 10.1016/j.seizure.2019.01.021 [PubMed: 30711397]
10. Chen S, He L, Huang A, Boehringer R, Robert V, Wintzer ME, et al. (2020). A hypothalamic novelty signal modulates hippocampal memory. Nature, 586(7828), 270–274. 10.1038/s41586-020-2771-1 [PubMed: 32999460]
11. Cleary AM, & Brown AS (2022). The Déjà vu Experience (2
12. Cleary AM, & Claxton AB (2018). Déjà vu: An illusion of prediction. Psychological Science, 29, 635–644. [PubMed: 29494276]
13. Cleary AM, Neisser J, McMahan T, Parsons TD, Alwaki A, Okada N, et al. (2021). Subjective distinguishability of seizure and non-seizure déjà vu: A case report, brief literature review, and research prospect. Epilepsy and Behavior, 125. 10.1016/j.yebeh.2021.108373
14. Corkin S, Amaral DG, Gonzalez RG, Johnson KA, & Hyman BT (1997). H.M.'s medial temporal lobe lesion: Findings from magnetic resonance imaging. Journal of Neuroscience, 17, 3964–3979. [PubMed: 9133414]
15. De Brigard F (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. Synthese, 191, 155–185. 10.1007/s11229-013-0247-7

16. Devereux G (1967). Fausse non-reconnaissance: Clinical sidelights on the role of possibility in science. Bulletin of the Menninger Clinic, 31, 69–78. [PubMed: 6039181]
17. Devitt AL, Addis DR, & Schacter DL (2017). Episodic and semantic content of memory and imagination: A multilevel analysis. Memory & Cognition, 45, 1078–1094. 10.3758/s13421-017-0716-1 [PubMed: 28547677]
18. Dokic J (2014). Feeling the past: A two-tiered account of episodic memory. Review of Philosophy and Psychology, 5, 413–426. 10.1007/s13164-014-0183-6
19. Donos C, Breier J, Friedman E, Rollo P, Johnson J, Moss L, et al. (2018). Laser ablation for mesial temporal lobe epilepsy: Surgical and cognitive outcomes with and without mesial temporal sclerosis. Epilepsia, 59, 1421–1432, 2018. [PubMed: 29893987]
20. Drane DL (2018). MRI-guided stereotactic laser ablation for epilepsy surgery: Promising preliminary results for cognitive outcome. Epilepsy Research, 142, 170–175. [PubMed: 28964596]
21. Dukes D, Abrams K, Adolphs R, Ahmed ME, Beatty A, Berridge KC, et al. (2021). The rise of affectivism. Nature Human Behaviour, 5(7), 816–820. 10.1038/s41562-021-01130-8
22. Gadian DG, Aicardi J, Watkins KE, Porter DA, Mishkin M, & Vargha-Khadem F (2000). Developmental amnesia associated with early hypoxic-ischaemic injury. Brain, 123, 499–507. 10.1093/BRAIN/123.3.499 [PubMed: 10686173]
23. Gardiner JM (2001). Episodic memory and autonoetic consciousness: A first-person approach. Philosophical Transactions of the Royal Society of London Series B Biological Sciences, 356(1413), 1351–1361. 10.1098/rstb.2001.0955 [PubMed: 11571027]
24. Gillinder L, Liegeois-Chauvel C, & Chauvel P (2022). What déjà vu and the "dreamy state" tell us about episodic memory networks. Clinical Neurophysiology, 136, 173–181. 10.1016/j.clinph.2022.01.126 [PubMed: 35189480]
25. Gloor P (1997). The temporal lobe and the limbic system. Oxford, UK: OUP.
26. Gorbach T, Pudas S, Bartrés-Faz D, Brandmaier AM, Düzel S, Henson RN, et al. (2020). Longitudinal association between hippocampus atrophy and episodic- memory decline in non-demented APOE ε4 carriers. Alzheimer's and Dementia, 12(1), Article e12110. 10.1002/dad2.12110
27. Hopkins RO, & Haaland KY (2004). Neuropsychological and neuropathological effects of anoxic and ischemic induced brain injury. Journal of the International Neuropsychological Society, 10, 957–961. 10.1017/s1355617704107108 [PubMed: 15803559]
28. Illman NA, Butler CR, Souchay C, & Moulin CJ (2012). Déjà experiences in temporal lobe epilepsy. Epilepsy research and treatment. , Article 539567. 10.1155/2012/539567, 2012. [PubMed: 22957231]
29. Isingrini M, Sacher M, Perrotin A, Taconnat L, Souchay C, Stoehr H, et al. (2016). Episodic feeling-of-knowing relies on noncriterial recollection and familiarity: Evidence using an online remember-know procedure. Consciousness and Cognition, 41, 31–40. 10.1016/j.concog.2016.01.011 [PubMed: 26849420]
30. Kant I (2008). Critique of pure reason. Muller, trans. Penguin classics, 1781 (Revised edition).
31. Kihlstrom JF (1995). Memory and consciousness: An appreciation of Claparède and recognition et moiitè. Consciousness and Cognition, 4(4), 379–386. 10.1006/ccog.1995.1045 [PubMed: 8750413]
32. Klein SB, & Nichols S (2012). Memory and the sense of personal identity. Mind, 121, 677–702.
33. Kripke S (1972). Naming and necessity. Cambridge, Mass: Harvard University Press.
34. Liscak R, Malikova H, Kalina M, Vojtech Z, Prochazka T, Murasic P, et al. (2010). Stereotactic radiofrequency amygdalohippocampectomy in the treatment of mesial temporal lobe epilepsy. Acta Neurochir (Wien), 152, 1291–1298. 10.1007/s00701-010-0637-2 [PubMed: 20361215]
35. Luo H, Zhao Q, Tian Z, Wu Z, Wang F, Lin H, et al. (2013). Bilateral stereotactic radiofrequency amygdalohippocampectomy for a patient with bilateral temporal lobe epilepsy. Epilepsia, 54, e155–e158. 10.1111/epi.12387 [PubMed: 24116395]
36. Mandler G (2008). Familiarity breeds attempts: A critical review of dual-process theories of recognition. Perspectives on Psychological Science, 3(5). https://escholarship.org/uc/item/0xh4x0j3.
37. Michaelian D, & Perrin. (2018). New directions in the philosophy of memory. Routledge Press.

38. Moulin CJA, Bell N, Turunen M, Baharin A, Akira R, & O'Connor. (2020). The the the the induction of jamais vu in the laboratory: Word alienation and semantic satiation. Memory. 10.1080/09658211.2020.1727519
39. Moulin C, Bell C, Turunen M, Baharin A, & O'Connor A (2021). The the the the induction of jamais vu in the laboratory: Word alienation and semantic satiation. Memory, 29(7), 933–942. 10.1080/09658211.2020.1727519 [PubMed: 32079491]
40. Moulin C, & Souchay C (2014). Epistemic feelings and memory. In Perfect TJ, & Lindsay D (Eds.), The SAGE handbook of applied memory (pp. 520–538). SAGE Publications Ltd. 10.4135/9781446294703.n29.
41. Ojemann GA, & Valiante T (2006). Resective surgery for temporal lobe epilepsy. In Miller JW, & Silbergeld DL (Eds.), Epilepsy surgery: Principles and controversy (pp. 403–413). New York: Taylor and Francis.
42. Park SA, Miller DS, Nili N, Ranganath C, & Boorman ED (2020). Map making: Constructing, combining, and inferring on abstract cognitive maps. Neuron, 107, 1226–1238. [PubMed: 32702288]
43. Pedersen NP, Ferrari L, Venner A, Wang JL, Abbott SBG, Vujovic N, et al. (2017). Supramammillary glutamate neurons are a key node of the arousal system. Nature Communications, 8, 1405.
44. Perrin D, Michaelian K, & Sant'Anna A (2020). The phenomenology of remembering is an epistemic feeling. Frontiers in Psychology, 11, 1531. 10.3389/fpsyg.2020.01531 [PubMed: 32719642]
45. Proust J (2013). The philosophy of metacognition. Oxford University Press.
46. Roy DS, Kitamura T, Okuyama T, Ogawa SK, Sun C, Obata Y, et al. (2017). Distinct neural circuits for the formation and retrieval of episodic memories. Cell, 170, 1000–1019. 10.1016/j.cell.2017.07.013 [PubMed: 28823555]
47. Rubin D (2022). A conceptual space for episodic and semantic memory. Memory & Cognition, 50(3), 464–477. 10.3758/s13421-021-01148-3 [PubMed: 33650021]
48. Russell B, & 2013. (1921). The analysis of mind. https://www.gutenberg.org/files/2529/2529-h/2529-h.htm.
49. Schacter DL, Addis DR, & Buckner RL (2007). Remembering the past to imagine the future: The prospective brain. Nature Reviews Neuroscience, 8, 657–661. 10.1038/nrn2213 [PubMed: 17700624]
50. Schwartz B (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. Psychonomic Bulletin & Review, 6(3), 379–393. [PubMed: 12198776]
51. Schwartz BL, & Metcalfe J (2011). Tip-of-the-tongue (TOT) states: Retrieval, behavior, and experience. Memory & Cognition, 39(5), 737–749. [PubMed: 21264637]
52. Scoville WB, & Milner B (1957). Loss of recent memory after bilateral hippocampal lesions. Journal of Neurology, Neurosurgery, and Psychiatry, 20, 11–21. [PubMed: 13406589]
53. Silbermann I (1963). The jamais phenomenon with reference to fragmentation. Psychoanalytic Quarterly, 32, 181–191. [PubMed: 13992833]
54. Sno HN (1994). A continuum of misidentification symptoms. Psychopathology, 27, 144–147. [PubMed: 7846229]
55. Sno HN (2000). D´ejà vu and jamais vu. In Berrios GE, & Hodges JR (Eds.), Memory disorders in psychiatric practice (pp. 338–347). Cambridge University Press. 10.1017/CBO9780511530197.016.
56. Szpunar KK, Spreng RN, & Schacter DL (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. Proceedings of the National Academy of Sciences, 18414–18421. 10.1073/pnas.1417144111
57. Tulving E (1972). Episodic and semantic memory. In Tulving E, & Donaldson W (Eds.), Organization of memory (pp. 381–403). New York: Academic Press.
58. Tulving E (1983). Elements of episodic memory. Oxford University Press.
59. Tulving E (1989). Memory: Performance, knowledge, and experience. European Journal of Cognitive Psychology, 1, 3–26.

60. Vargha-Khadem F, Gadian DG, Watkins KE, Connelly A, Van Paesschen W, & Mishkin M (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. Science, 277, 376–380. [PubMed: 9219696]

61. Warren-Gash C, & Zeman A (2014). Is there anything distinctive about epileptic d´ejàvu? Journal of Neurology, Neurosurgery, and Psychiatry, 85, 143–147. [PubMed: 23315620]

62. Williams D (1953). The elements of being. The Review of Metaphysics, 7(1), 3–18.

63. Willie JT, Laxpati NG, Drane DL, Gowda A, Appin C, Hao C, et al. (2014). Real-time magnetic resonance-guided stereotactic laser amygdalophippocampotomy (SLAH) for mesial temporal lobe epilepsy. Neurosurgery, 74, 569–585. [PubMed: 24618797]

64. Zola-Morgan S, Squire LR, & Ramus SJ (1994). Severity of memory impairment in monkeys as a function of locus and extent of damage within the medial temporal lobe memory system. Hippocampus, 4, 483–495. [PubMed: 7874239]

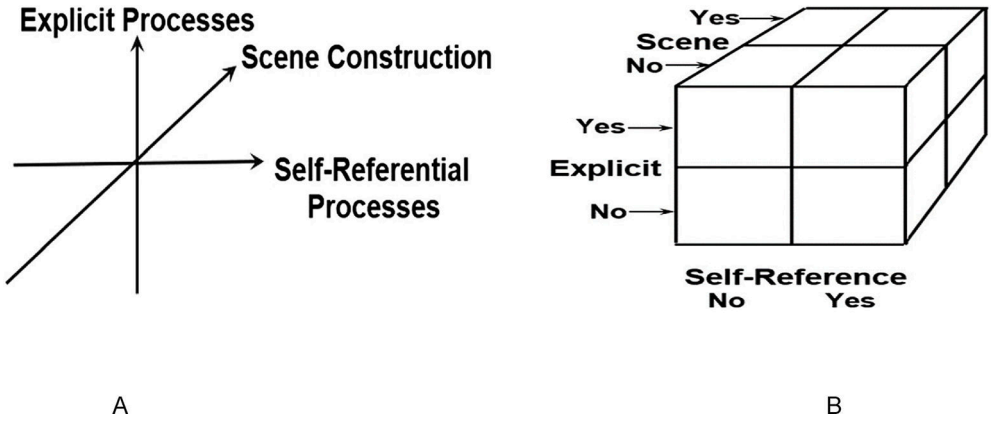

Fig. 1.
AB (Rubin, 2022).

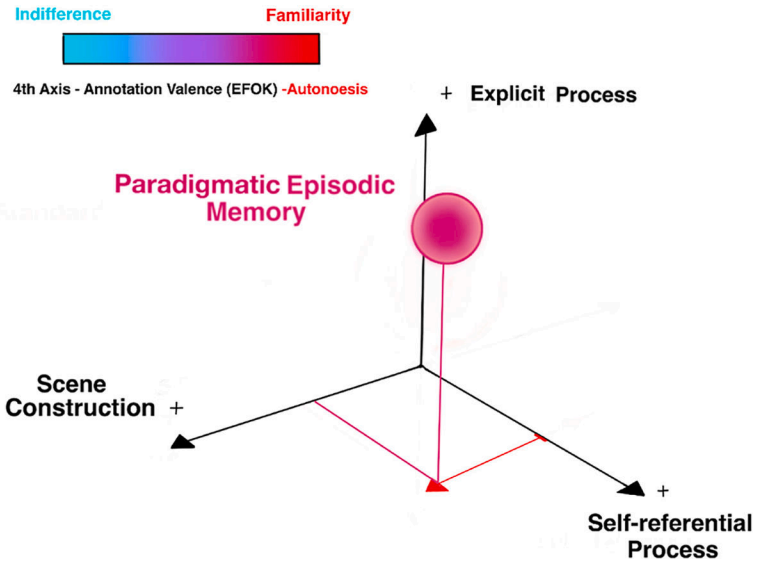

Fig. 2.
A typical episodic memory state, plotted in four dimensions.

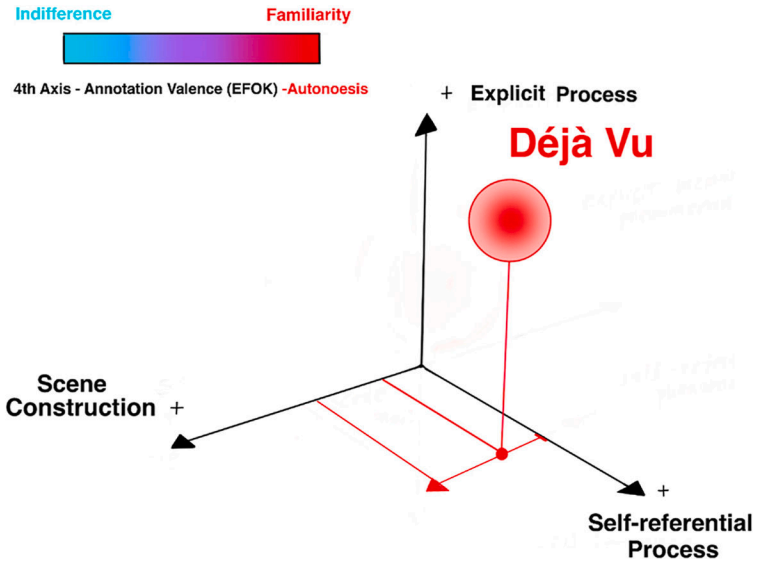

Fig. 3.
A state of déjà vu, plotted in four dimensions.

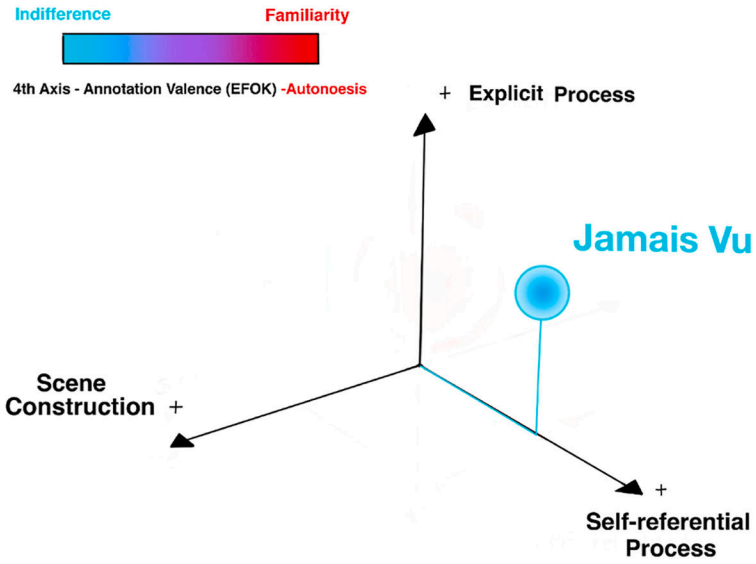

Fig. 4.

Jamais vu as evoked in a semantic satiation task, plotted in fours dimensions.

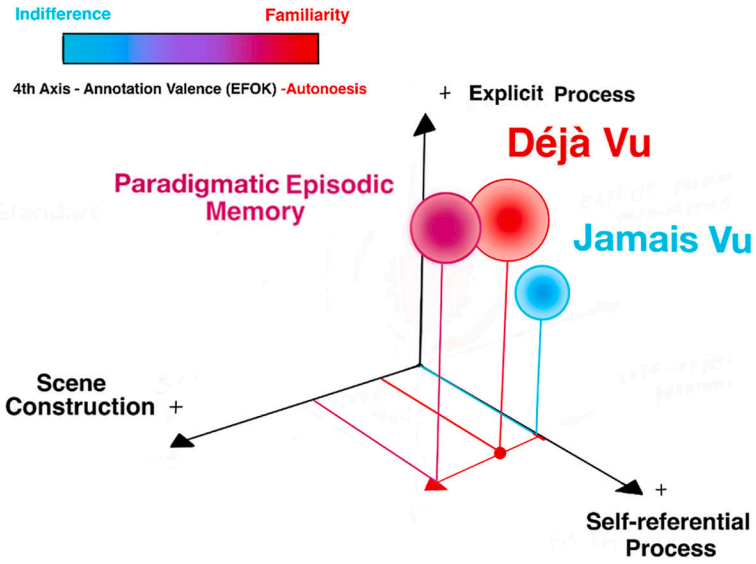

Fig. 5.
The three variants of Mental Time Travel, plotted together in four dimensions.

Table 1.

Classical assumption of "concordance" between memory contents and experiences.

Empty Cell	**Noetic Phenomenology FOK, TOT**	**Autonoetic Phenomenology FOP, EFOK**
Semantic Memory	"Just knowing" in concordance with metamemory feelings about one's knowledge	Empty Set No first-order content and no corresponding metamemory feeling
Episodic Memory	Empty Set No first-order content and no corresponding metamemory feeling	explicit scene memory and self-reference in concordance with the feeling of mental time travel

Table 2.

Memory contents and experiences, without the assumption of “concordance”.

Empty Cell	**Noetic Phenomenology FOK, TOT**	**Autonoetic Phenomenology FOP or EFOK**
Semantic Memory	Paradigmatic “just knowing” Memory for facts & propositions, with accompanying metamemory feelings	**Autonoetic Knowing** déjà vu as “just reliving” i.e. the odd feeling of mental time travel but without memory content for where & when
Episodic Memory	**Dis-autonoesis** Presence of episodic content but lack or distortion of the feeling of mental time travel and/or familiarity Jamais vu, RB	Paradigmatic “mental time travel” Explicit scene memory with self-reference and EFOK