



Target Article

*The authors contributed equally to this paper and share first authorship.

Cite this article: Barzykowski K, Moulin CJA. (2023) Are involuntary autobiographical memory and déjà vu natural products of memory retrieval? *Behavioral and Brain Sciences* **46**, e356: 1–67. doi:10.1017/S0140525X22002035

Target Article Accepted: 5 September 2022
Target Article Manuscript Online: 16 September 2022
Commentaries Accepted: 11 December 2022

Keywords:

autobiographical memory; déjà vu; involuntary cognitions; involuntary memories; retrieval processes; spontaneous cognition; threshold model

What is Open Peer Commentary? What follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (p. 22) and an Authors' Response (p. 59). See [bbsonline.org](https://www.cambridge.org/bbs/online) for more information.

^aApplied Memory Research Laboratory, Institute of Psychology, Faculty of Philosophy, Jagiellonian University, Kraków, Poland; ^bLaboratoire de Psychologie et Neurocognition, Université Grenoble Alpes, Grenoble, France and ^cInstitut Universitaire de France
krystian.barzykowski@uj.edu.pl

Abstract

Involuntary autobiographical memories (IAMs) and déjà vu are phenomena that occur spontaneously in daily life. IAMs are recollections of the personal past, whereas déjà vu is defined as an experience in which the person feels familiarity at the same time as knowing that the familiarity is false. We present and discuss the idea that both IAMs and déjà vu can be explained as natural phenomena resulting from memory processing and, importantly, are both based on the same memory retrieval processes. Briefly, we hypothesise that both can be described as “involuntary” or spontaneous cognitions, where IAMs deliver content and déjà vu delivers only the feeling of retrieval. We map out the similarities and differences between the two, making a theoretical and neuroscientific account for their integration into models of memory retrieval and how the autobiographical memory literature can explain these quirks of daily life and unusual but meaningful phenomena. We explain the emergence of the déjà vu phenomenon by relating it to well-known mechanisms of autobiographical memory retrieval, concluding that IAMs and déjà vu lie on a continuum.

1. Introduction

It's exactly the same as the last time, it's all come back to me ... it's behind, it's always thoughts from childhood, it's always visual, it's a place behind the house, the field where my father put his car, near a lake ... It's not always the same countryside; I've forgotten the story of this countryside ... Yes, it's pleasant because we were going to get the car from behind the house, it's a happy memory, it's never unpleasant.

Vignal, Maillard, McGonigal, and Chauvel (2007, p. 92)

I have the impression that everything around me has been here before, that it has already happened, I feel as if I am going backwards in time.

Vignal et al. (2007, p. 92)

I could be carrying out an everyday activity – e.g., driving/attending lecture – and visual images appear in my head – possibly of people I know/places I've been (with no relevance to what's going on in reality). I could be doing the housework and certain melodies I hadn't heard for a while pop into my head.

Kvavilashvili and Mandler (2004, p. 89)

It feels like you're living an experience that you've already lived through or perhaps dreamed. As each small detail of the experience unfolds it seems as though you knew it was going to happen--as if you could have described the scene before it even existed. Each time I feel a strong conviction that I've seen all of it happen before.

Jersakova, O'Connor, and Moulin (2015, data available online)

These quotes capture two different spontaneous cognitions, in two different populations. The first two refer to reports of spontaneous thoughts cued by electrical stimulation of the brain (for a review, see Curot et al., 2017). The second two are subjective reports of the same phenomena in healthy people. One phenomenon concerns the reproduction of a fully formed scene from the personal past, an involuntary autobiographical memory (IAM; Berntsen, 1996, 2010), the other concerns a vaguer feeling described as déjà vu (Moulin, 2018). Here, we outline how these phenomena are related, with the view that theories of memory function should take them into account.

The starting point for our synthesis comes from Penfield's (e.g., Penfield & Perot, 1963) intracortical stimulation studies, whereby current applied to the temporal cortex produced “interpretative illusions.” In their review of 520 stimulations of patients' temporal lobes, Penfield and Perot (1963) reported that 7% of stimulations produced “responses” devoid of content, and 10% of stimulations produced “hallucinations,” which included sensory content, and which were mostly autobiographical experiences. In contemporary work, Vignal et al. (2007) showed that stimulation of the same electrical contact in the temporal lobe either

produced déjà vu (with lower voltages and more restricted discharges) or a more fully formed specific memory. It seems that at least some forms of retrieval from the personal past and déjà vu share underlying neurological and/or cognitive processes, and the difference between them may be expressed on a continuum. Where sufficient activation exists from a cue in the environment, there is the experience of effortless retrieval, whereas where retrieval fails it can prompt unusual and seemingly out-of-context feelings. These experiences are not restricted to electrical stimulation; involuntary memories and déjà vu occur in healthy participants, and can be provoked in the laboratory.

We map out the similarities and differences between these phenomena (see Table 1), and provide an account for their integration into theories of memory retrieval. Both can be described as “involuntary” or “spontaneous cognitions,” one delivering content, and the other delivering only a feeling. The development of a theory which integrates these spontaneous cognitions into a memory retrieval framework is thus of value, especially since a novel development in the field is that such phenomena are underpinned by the default mode network (DMN; e.g., Andrews-Hanna, Reidler, Huang, & Buckner, 2010; Buckner, Andrews-Hanna, & Schacter, 2008; Raichle, 2015; Stawarczyk, Majerus, Maquet, & D’Argembeau, 2011; Van Calster, D’Argembeau, Salmon, Peters, & Majerus, 2017), and changes in spontaneous cognition might be one of the earliest signs of temporal-lobe pathology (Kvavilashvili, Niedźwieńska, Gilbert, & Markostamou, 2020). Any account of memory retrieval ought to consider these two phenomena.

There has also been a shift of attention in cognitive psychology from studying goal-oriented, voluntary, and effortful forms of human cognition to studying phenomena that come to mind spontaneously and/or without any preceding intention to think about them. This has resulted in the emergence of several independent research fields on (1) mind wandering and daydreaming (e.g., Christoff, Irving, Fox, Spreng, & Andrews-Hanna, 2016; Smallwood & Schooler, 2015), (2) IAM (e.g., Berntsen, 2009; Mace, 2007) and automatically retrieved voluntary memories (Barzykowski & Staugaard, 2016; Uzer, Lee, & Brown, 2012), (3) spontaneous future thinking (Cole & Kvavilashvili, 2019, 2021), (4) spontaneous retrieval of prospective memory tasks

(Gilbert, Hadjipavlou, & Raelison, 2013; Kvavilashvili & Rummel, 2020; Niedźwieńska & Barzykowski, 2012; Niedźwieńska, Rendell, Barzykowski, & Leszczyńska, 2014), (5) involuntary semantic memories or mind-pops (Kvavilashvili & Mandler, 2004), (6) intrusive memories (Marks, Franklin, & Zoellner, 2018), (7) involuntary musical imagery (earworms; Hyman et al., 2013), (8) déjà vu and associated spontaneous metacognitions (Moulin, 2018), and (9) game transfer phenomena (Ortiz De Gortari & Griffiths, 2015). As yet, these spontaneous phenomena from daily life are studied separately (but see Barzykowski, Hajdas, Radel, Niedźwieńska, & Kvavilashvili, 2021a, 2022; Barzykowski, Radel, Niedźwieńska, & Kvavilashvili, 2019b; Plimpton, Patel, & Kvavilashvili, 2015; Vannucci, Pelagatti, Hanczakowski, & Chiorri, 2019), significantly impeding our understanding of their nature and shared underlying mechanisms. To move towards an integration of these phenomena, we start with IAMs and déjà vu; specific forms of spontaneous cognition for the personal past, where there have been significant developments recently in the literature.

1.1. Involuntary autobiographical memories (IAMs)

Autobiographical memory is the ability to remember events from our personal past (Conway & Pleydell-Pearce, 2000). Theories distinguish between involuntary and voluntary retrieval in autobiographical memory as a consequence of conscious intention (i.e., wanting to remember). IAMs are the reminiscence of personal events that come to mind without any conscious attempt at retrieval (Berntsen, 1996, 2010). They are contrasted with voluntary autobiographical memories (e.g., Barzykowski & Staugaard, 2016, 2018; Schlagman & Kvavilashvili, 2008). Voluntary memories typically involve an effortful and time-consuming strategic search (Botzung, Denkova, Ciuciu, Scheiber, & Manning, 2008; Burgess & Shallice, 1996; Conway, 1990, 2005; Conway & Loveday, 2010; Conway & Pleydell-Pearce, 2000), but they may also be directly recalled (i.e., automatically and effortlessly) in a similar, yet still somewhat different, way as involuntary memories (Uzer et al., 2012; see also Barzykowski & Staugaard, 2016; Harris, O’Connor, & Sutton, 2015; Uzer & Brown, 2017). When we use our memory intentionally, for example, when trying to recall whether we have already paid a conference fee, we retrieve voluntary autobiographical memories. However, memories may sometimes pop into our mind without any retrieval attempt, for example, when driving, a memory of meeting our partner for the first time may enter into consciousness without being sought-for.

Involuntary memories are now treated as a common example of remembering (e.g., Berntsen, 2010, 2015; Brewin, Gregory, Lipton, & Burgess, 2010; Clark, Mackay, & Holmes, 2013; Moulds & Krans, 2015) and as a phenomenon worthy of investigating in and outside the laboratory (e.g., Berntsen, 1996; Roberts, McGinnis, & Bladt, 1994; Schlagman & Kvavilashvili, 2008). IAMs were overlooked in cognitive psychology for several decades (e.g., Miller, 1962/1974), but starting from the nineties (e.g., Berntsen, 1996; Roberts et al., 1994), they have been investigated using three main methods (see Berntsen, 2009, for a more detailed review): Survey methods (e.g., Berntsen, Rubin, & Salgado, 2015), structured diaries (e.g., Johannessen & Berntsen, 2010; Mace, 2005), and experimental methods (e.g., Ball, 2007; Barzykowski & Niedźwieńska, 2016, 2018a; Mace, 2006; Schlagman & Kvavilashvili, 2008; Vannucci, Batool, Pelagatti, & Mazzoni, 2014).

KRYSTIAN BARZYKOWSKI is an autobiographical memory specialist in experimental cognitive psychology based at the Institute of Psychology at Jagiellonian University (Kraków, Poland). Over the years, he has become an expert on involuntary and voluntary memory studies, which has resulted in several peer-reviewed papers on involuntary and voluntary memory retrieval. He is a recipient of several research grants (e.g., from National Science Center, Poland, Polish National Agency for Academic Research) and has earned several scholarships and awards for his research and scientific achievements.

CHRIS J. A. MOULIN has conducted research on the cognitive neuropsychology of memory since his PhD at the University of Bristol. Currently, he is a professor of cognitive psychology at the Université Grenoble Alpes in France. He is a laureate of the Institut Universitaire de France. An author on several peer-reviewed articles on autobiographical memory, and an expert on the subjective experience of memory and metamemory, he recently published a monograph, *The Cognitive Neuropsychology of Déjà Vu* for the essays in cognitive psychology series (Psychology Press/Routledge).

Table 1 (Barzykowski and Moulin). The summary of similarities and differences between déjà vu and involuntary autobiographical memories (partially inspired by Kvavilashvili, 2015, Figure 1).

	Déjà vu	Involuntary autobiographical memories
Basic definition	Feeling of familiarity combined with the knowledge that the feeling is false	Recollections of personal events that come to mind without any conscious attempt at retrieval
Key characteristics	<ul style="list-style-type: none"> • Automatic: Effortless • Cannot be willfully generated • Unexpected • Infrequent • The feeling of familiarity with one's personal past 	<ul style="list-style-type: none"> – Automatic: Effortless – Involuntary: No preceding conscious attempt – Unexpected – Direct relation to one's personal past
Similarities/shared properties		
Population	Normal and clinical	
Retrieval	Involuntary/spontaneous	
Effort	Automatic: None or minimal	
Appearance	Incidental: Accidentally triggered/arousing	
Feeling of retrieval	Mostly explicitly experienced	
Retrieval latency	Low: Quickly retrieved	
Methods used	Questionnaire, experimental, diary, neuropsychological	
Relevance to personal past	Grounded in/immersed in/related to personal past	
Distress	Rarely: None to moderate	
Differences		
Manifestation	Elusive sensation of false familiarity	Memory of personal events, experiences
Memory content	Absence of content: No direct relation to given past experience and details	Usually direct relation to given past experience with different levels of details
Feeling of familiarity	Central, overwhelming, explicit	Peripheral, secondary, extremely brief or non-existent
Time orientation	Feeling of pastness for present moment	The recollection of past events/experiences
Cognitive demand	Highly cognitively demanding: Only experienced by people with sufficient cognitive resources to oppose and detect the false feeling of familiarity	Only mildly cognitively demanding (although limited by cognitive load)
Accuracy	False familiarity Knowledge that familiarity is inaccurate	Rather accurate: Knowing/believing that what is retrieved is correct and true
Prevalence	Relatively Low/infrequent/rare	Relatively high/frequent/common
Recurrence	Unknown – presumably low – ironically, one does not get the same déjà vu often	Low to high – one does not get the same involuntary memory very often but some memories may become recurrent and repeated
Cue	Configural, contextual similarity (Gestalt account), and/or perceptual prime (in experimental work)	External or internal identifiable cues
Source	Unknown or “unknown past”: Source amnesia	Rather easily identifiable external/internal cues/past events
Accessibility	Rather high: Easily passes awareness threshold	High: Strong and intense experience – easily passes awareness threshold
Phenomenological characteristics (e.g., vividness, clarity)	Mostly described as intense – but no explicit content	Mostly sound and intense: Identifiable phenomenological characteristics
Valence	Pleasant experience typically – but difficult to compare since no content; at the same time may be felt as disturbing and unsettling due to it's incomprehensible nature	Mostly positive; also neutral and negative
Avoidance	Unknown	None or minimal – differentiates them from recurrent or traumatic memories
Disruption to current cognitive processing	Unknown	None or minimal
Metacognition	Metacognitive: Top-down Higher order interpretation of retrieval processes, feeling of familiarity	Minimally metacognitive – Bottom-up: Triggered/memory's phenomenological characteristics draw one's memory-related attention

(Continued)

Table 1 (Continued.)

	Déjà vu	Involuntary autobiographical memories
Self-awareness	High (always present)	High to none (one may not be aware of having IAMs)
The role of attributions (e.g., beliefs in retrieval)	High	Rather non to minimal
Functionality	Error of the memory metacognitive/metamemory illusion/error; side effect of typical involuntary processing of contextual/environmental cues	While they may be side effect of typical involuntary processing of contextual/environmental cues, IAMs may serve several functions and roles
Neurological basis	Familiarity signal produced by medial temporal-lobe circuitry, upstream conflict detection in cingulate cortex/prefrontal cortex	Posterior cingulate cortex, dorsolateral prefrontal cortex (inactivated), posterior parietal regions (e.g., inferior parietal lobule), medial temporal-lobe areas, increased activation in the hippocampus, precuneus, and ventrolateral prefrontal cortex

The laboratory method developed by Schlagman and Kvavilashvili (2008; with further modifications, e.g., Barzykowski, Niedźwieńska, & Mazzoni, 2019a; Plimpton et al., 2015; Vannucci et al., 2019) allows the examination of the retrieval phase of this spontaneous phenomenon under controlled conditions (e.g., Barzykowski, 2014; Barzykowski & Niedźwieńska, 2012). Participants are engaged in a monotonous vigilance task requiring little attention (such as detecting 15 infrequent target vertical lines in a stream of 785 slides with horizontal lines) while exposed to irrelevant word phrases (e.g., *buying a baguette, a wonderful smile, an unpleasant conversation*), some of which may incidentally trigger involuntary memories (e.g., *I remember buying fresh and still warm bread from the bakery's back window when getting back from a crazy all night party; this was something I very much needed when going home at 4am*). Participants are instructed to write down any spontaneously occurring thought and/or memories during the vigilance task. Schlagman and Kvavilashvili's (2008) participants reported on average approximately six involuntary memories (ranging from 1 to 24) during laboratory sessions.

While IAMs were treated at first as a relatively rare phenomenon (e.g., Davachi & Dobbins, 2008), they are now considered common in daily life (Berntsen, 2010, 2015; Brewin et al., 2010; Clark et al., 2013; Moulds & Krasn, 2015). Involuntary remembering may be even more frequently experienced in daily life than voluntary and effortful autobiographical retrieval (e.g., Rasmussen & Berntsen, 2011; Uzer et al., 2012). The high prevalence of IAMs is also demonstrated by people's tendency to classify automatically (i.e., effortlessly) retrieved voluntary memories as involuntary memories even though these memories were retrieved in response to a direct instruction to do so (e.g., Barzykowski, Staugaard, & Mazzoni, 2021b; Sanson, Cardwell, Rasmussen, & Garry, 2020). IAMs may be considered as a universal well-known and frequently experienced phenomenon that commonly occurs in the general population and which occurs in healthy autobiographical memory.¹ While typical frequency estimates vary depending on the method used (e.g., Laughland & Kvavilashvili, 2018), the average frequency of IAMs is in the order of a dozen per day (e.g., roughly 20 memories per day; Rasmussen & Berntsen, 2011).

The results of naturalistic diary studies, where participants are instructed to record and describe IAMs and their accompanying context immediately as they occur in everyday life (e.g., Berntsen, 1996; Laughland & Kvavilashvili, 2018), have shown that IAMs are more likely to be reported when attention is diffuse (Berntsen, 1996; 2009), and when the individual is engaged in an

automatic activity with low attention and low cognitive demands (e.g., washing-up, walking, ironing). They also arise in response to incidental external and/or internal cues that usually overlap with key features of the memory content (e.g., seeing a given mug on a desk may trigger a certain past episode of receiving it as a gift; Berntsen, 1998). This is supported by data from laboratory settings: The content of IAMs directly corresponds to experimentally provided cues (e.g., Barzykowski & Niedźwieńska, 2018b; Plimpton et al., 2015; Vannucci, Pelagatti, Hanczakowski, Mazzoni, & Paccani, 2015, 2019). This idea, that IAMs are in fact a response to cues in the environment even though they are experienced as involuntary, is a critical aspect of our argument and we explain this focussing on the notion of cognitive control.

It could be argued that the key difference between involuntary and voluntary memories is the presence or not of cognitive control. When trying to deliberately retrieve a given memory (e.g., *eating Ramen with close friends two weeks ago*) in response to a certain cue (e.g., *eating Ramen*), one has to actively and strategically search for a representation that specifically relates to that cue. Thus, there is a cascade of cognitive processes leading to successful memory retrieval: (1) Actively elaborating on the cue, (2) monitoring the stream of consciousness looking for a thought corresponding to a cue, (3) deciding whether the thought relates to the searched episode from the personal past and (4) deciding to terminate search or continue the search until a better match is found (see also Mace, Clevinger, Delaney, Mendez, & Simpson, 2017, for mental strategies used in voluntary recall).

Involuntary retrieval circumvents such search processes as there are no pre-set requirements to be met and the involvement of cognitive control in the memory retrieval is minimised. The presence of cognitive control is also reflected by longer retrieval latencies for voluntary than involuntary memories (e.g., Barzykowski et al., 2021b; Schlagman & Kvavilashvili, 2008). In addition, Barzykowski et al. (2021b) recently also reported data suggesting that voluntary retrieval entails not only increased cognitive control but also more awareness during memory retrieval, which allows participants to retrospectively reflect upon and evaluate their retrieval. Involuntary memories thus come to mind rapidly, and thus somewhat devoid of phenomenology, since there is no feeling of effort.

Whilst there are only a few neuroscientific studies which help elucidate the nature of the difference between voluntary and involuntary memories, there is a consensus view on the neuroscience of cognitive control more generally. Traditionally, the frontal lobes are seen as a critical structure in a network of regions

(essentially a frontal–cingulate–parietal network) responsible for cognitive control (e.g., Cai et al., 2016). This network is involved in bringing to attention task-relevant information, distributing attentional resources and suppressing inappropriate behavioural responses. In a review, Menon and D’Esposito (2022) outline six different large-scale functional networks involved in cognitive control, proposing that cognitive control is implemented by dynamic interactions among prefrontal cortex (PFC) networks embedded in a “global brain architecture.” In memory retrieval more specifically, Eichenbaum (2017) proposes that during recall, contextual cues are processed by the ventral hippocampus, which sends this information to the prefrontal cortex, which in turn “biases” the retrieval of context-appropriate memories from the dorsal hippocampus.

These neuroscientific descriptions of the complex, dynamic interactions between the prefrontal cortex and the temporal lobe are described without reference to comparison between voluntary and involuntary retrieval, and so it is difficult to know the implication of the prefrontal cortex in involuntary retrieval, despite causal mechanisms being identified in the literature (i.e., prefrontal activation occurring before, or leading to, cognitive control and memory retrieval [Cai et al., 2016; Eichenbaum, 2017]). In their Trace Transformation Theory, Sekeres, Winocur, and Moscovitch (2018) propose a similar interaction between the same structures, but also specify difference between the anterior and posterior hippocampus in humans, from more coarse representations in the anterior hippocampus and more fine-scaled representations in the posterior hippocampus (e.g., Brunec et al., 2018).

The neuroscientific literature describes a network with a to-and-fro between retrieval and control mechanisms driven by cues in the environment, which inspires our view here. But note that most neuroscientific accounts of this process do not differentiate between voluntary and involuntary retrieval (perhaps because these are more subtle differences in the non-human brain and are difficult to operationalise experimentally). However, cognitive control is a key component of neuroscientific accounts of retrieval and our focus here is what occurs cognitively when memories are involuntarily retrieved: Does this imply less control, and possibly less of a causal activation derived from the prefrontal cortex?

To examine this, Hall et al. (2014) compared voluntary and involuntary retrieval, finding that voluntary retrieval elicited greater activity in dorsal frontal regions while keeping other sub-components of the retrieval network (e.g., medial temporal, ventral occipitotemporal, and ventral parietal regions) similarly engaged. Previous studies (e.g., Hall, Gjedde, & Kupers, 2008; Hall et al., 2014; Kompus, Eichele, Hugdahl, & Nyberg, 2011; for review, see Kvavilashvili et al., 2020) suggest important roles for the posterior cingulate cortex (involved during both involuntary and voluntary retrieval), dorsolateral prefrontal cortex (inactivated during involuntary but activated during voluntary retrieval), posterior parietal regions (e.g., inferior parietal lobule),² and medial temporal-lobe areas in the involuntary retrieval processes. Therefore, involuntary retrieval compared to voluntary retrieval may not need additional engagement of the prefrontal cortex.

The phenomenology of IAMs is of them being spontaneous and this is corroborated by the neuroimaging findings that converge on a lack of cognitive control in involuntary retrieval. In this way, they might be thought of as part of a family of “stimulus unrelated thoughts” including mind wandering, which is

described as being thought decoupled from current perceptual inputs (e.g., Mills, Zamani, White, & Christoff, 2021; Schooler et al., 2011). However, as stated above, IAMs are in fact often related to easily identifiable cues in the environment (e.g., Plimpton et al., 2015; Schlagman & Kvavilashvili, 2008). Mace (2004) showed that in everyday life, IAMs were most likely to be triggered by abstract cues (e.g., thoughts, language-based referents to the original past event) rather than sensory/perceptual cues (e.g., raw sensory experiences such as taste or smell). This finding suggests that although sensory and perceptual experiences may sporadically trigger an involuntary memory, more often than not they are evoked by objects, scenes, themes, abstract linguistic-based cues, and activation of autobiographical memory schemas. Several authors stress the relationship between mind wandering and involuntary memories (see Cole & Kvavilashvili, 2019; Plimpton et al., 2015), stating that retrieving an IAM is not related to the ongoing task, but it is related to information in the environment. To summarise, we propose that retrieval in involuntary memory is relatively rapid and although cued by the environment, arrives without an awareness of memory search, and is as such a relatively pure, content-rich type of retrieval, as indicated in Figure 1.

1.2. Déjà vu

Déjà vu is a brief experience in which the person experiences false familiarity at the same time as knowing that the familiarity is false. A reasonable estimate is that about 80–90% of the population experience it at least once, and no more than 10 times a year (see Brown, 2003). It is therefore experienced far less frequently than IAMs. Questionnaire studies have examined the possible triggers of déjà vu, and although it appears to be more likely for scenes, conversations, and for familiar places and people, no real patterns emerge (see Moulin, 2018, for a review). It is not the case that déjà vu has a uniquely visual trigger: Even people who are blind report experiencing déjà vu³ (O’Connor & Moulin, 2006). The current conceptualisation is that déjà vu arises from an inappropriate activation of familiarity: It feels as if a memory is being accessed, but no such representation from the personal past is present (Brown, 2004). This idea that déjà vu might result when the relevant memory cannot be retrieved has inspired the paradigms developed for studying déjà vu.

Definitions converge on the idea of a clash in mental evaluations and a conflict in appraisals (e.g., Urquhart, Sivakumaran, Macfarlane, & O’Connor, 2021). For instance, one may experience déjà vu when entering a flat, having a strong feeling of having been in this flat before, although at the same time knowing that this is actually the first time you have visited your friend’s place. The déjà vu experience is thus also metacognitive: It arises out of a higher order interpretation of retrieval processes; it is not possible to have a déjà vu experience and not be aware of it. This critical feature distinguishes it from a false memory.

Whilst some authors have suggested the conflict in déjà vu arises from false familiarity which is top–down, or neurological in origin (the decoupled familiarity hypothesis; Illman, Butler, Souchay, & Moulin, 2012). This sees déjà vu as essentially an infrequent or random experience caused by a neurophysiological event, not unlike the electrical stimulations which start this article. Others, however, have suggested it is cued by stimuli which present some undetected conceptual or perceptual overlap with stored representations (e.g., the Gestalt Similarity account; Cleary et al., 2012). Déjà vu is a spontaneous mental event, in

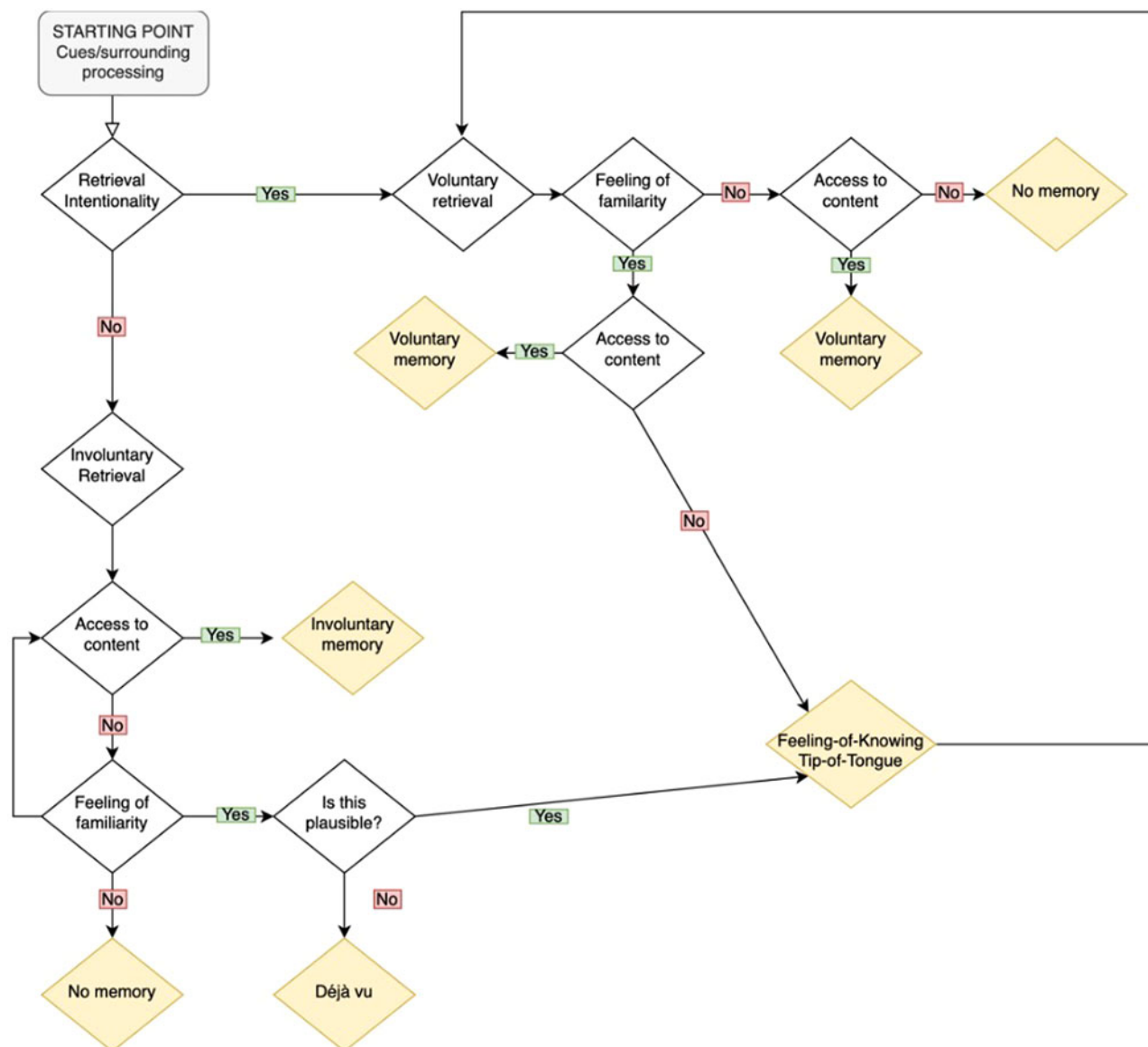


Figure 1. Flow chart representing the memory phenomena (autobiographical memory, déjà vu, tip of the tongue, feeling of knowing), as a result of (1) retrieval intentionality (involuntary vs. voluntary), (2) memory content accessibility (accessible/inaccessible), and (3) feeling of familiarity (present/absent). Outcomes of the retrieval process are: *No memory* (whereby nothing is retrieved that is experienced as a memory), *déjà vu*, *involuntary memory*, *voluntary memory*, and the *feeling of knowing/tip-of-the-tongue*, this latter, which is evaluated as plausible despite the lack of retrieved content is experienced as a frustrating sense of familiarity for a currently unretrieved representation. It is a sensation which provokes a search in memory, hence its link to voluntary retrieval process. *Access to content*: Access to content implies complete recollective retrieval of the personal past including the phenomenology of remembering and a sense of successful retrieval. *Feeling of familiarity*: Feeling of familiarity implies a subjective experience of fluency devoid of any content. *Voluntary and involuntary retrieval*: These labels refer to generic memory processes which are either will fully engaged (e.g., memory search, cue elaboration, generation of associations) or which are provoked by cues in the environment.

that it is not directly related to the current goals of conscious processing, and it is not sought for. However, as with IAMs, it can be triggered by cues in the environment.

The idea is that, for example, when entering the friend's flat for the first time, the configuration of the window, fireplace, and sofa match a stored representation, and give a certain configural fluency (e.g., Cleary et al., 2012; Cleary & Claxton, 2018), such that a feeling of familiarity for the environment is generated which relates strictly to one's personal past, even though it is known that this is the first visit to the apartment. The room feels familiar because it *reminds us* of something we have already encountered in our past, but we are not able to retrieve the source

of the familiarity: We are unaware of the resemblance. Similar theories have been proposed since the nineteenth century (e.g., Knight, 1895), and may involve anodyne processes such as reading a description of a place and then visiting the same place or somewhere similar (e.g., Hawthorne, 1863).

Such experiences can be provoked in the laboratory. In the first of a series of influential experiments, Brown and Marsh (2008, see also 2009, 2010) used unattended processing to change beliefs about memory and prior occurrence, using a superficially presented stimuli to produce a déjà vu-like experience. This would be comparable with having a quick glance at a scene, and then subsequently processing the information more fully – what has

been described as a cognitive “double take.” Critically, the same information is *actually* processed twice, which is presumably not the case for our example in our friend’s flat. Cleary and colleagues (e.g., Cleary, Ryals, & Nomi, 2009) sought to explain déjà vu, not in terms of an exactly matching unattended processing of a stimulus, but as a perceptual overlap, basing their empirical work on the recognition without identification paradigm (see Cleary, 2008). Their hypothesis was that déjà vu was provoked by configural overlap in perception.

In their first experiment, they presented line drawings of scenes, such as an airstrip with a plane coming into land. This image is matched with second “test” image which shares the same general layout and elements, a picture of a pond. Instead of having the aeroplane above the horizon in the centre, there is a dragonfly, and the converging lines of the airstrip tapering into the distance are replaced by a similar pattern on the surface of the water. This gives two unique scenes which share a large perceptual similarity. Participants study the scenes with labels (e.g., *airstrip*) and at test they see the novel configurally similar scenes, which they cannot name (e.g., *pond*), since they were never studied. The configural similarity generates a feeling of familiarity, but when combined with the failure to produce a name (i.e., recognition without identification), it generates déjà vu. Subsequently, using virtual reality, Cleary et al. (2012) used conceptually the same method but used rooms and scenes (exactly like the flat example above), again producing déjà vu in their participants.

Drawing on this virtual reality approach, and using navigational paths through a virtual environment, Cleary and colleagues have mapped out the range of attributions and subjective experiences that accompany déjà vu (e.g., Cleary & Claxton, 2018; Cleary, Huebert, McNeely-White, & Spahr, 2019; Cleary, McNeely-White, Huebert, & Claxton, 2021a). In these experiments, Cleary and colleagues sought to explain the feeling of “prescience” in déjà vu – the ability to be able to predict the future. In the first work in this series, Cleary and Claxton showed that a feeling of familiarity with a turn in a path generated feelings akin to déjà vu, replicating their earlier work, but this feeling did not lead to above-chance ability to predict the next turn. Critically, however, déjà vu was accompanied by increased feelings of knowing the direction of the next turn, leading Cleary and colleagues to describe déjà vu as an illusion of prediction. In a follow up, Cleary, McNeely-White, Huebert and Claxton showed that prescience was not a ubiquitous experience in déjà vu, but that it was related to the level of familiarity felt; the more intense the phenomenology of familiarity, the more the participant felt like they could predict what was coming next. Cleary, Huebert, McNeely-White and Spahr went one stage further, asking participants to report on how things unfolded after the feeling of déjà vu, showing a “postdictive bias” as well as the illusion of prediction. Participants were more likely to say that the novel route unfolded in a way that they remember from the déjà vu: Something which, as in the earlier task, was not borne out in their ability to identify the right path.

Most recently, Huebert, McNeely-White, and Cleary (2022) tested the idea that feelings of familiarity in cued recall might provoke illusory feelings of recollection – being able to recall contextual specifics from a past event, in line with the observation of recollective confabulation patients, who have been described as having permanent déjà vu (Moulin, 2013). Again using the recognition without recall paradigm, Huebert et al. manipulated the amount of cue-target feature overlap. Increasing familiarity led to increased confidence in knowing a contextual detail of some

other information that was presented at study. In sum, using various iterations of the recognition without identification paradigm, Cleary and colleagues have repeatedly provoked feelings of familiarity in the laboratory which are described as déjà vu by participants. Crucially for our argument below, this feeling of familiarity has implications for illusions of prediction and recollection.

In a different experimental framework, Urquhart and colleagues (Urquhart & O’Connor, 2014; Urquhart et al., 2021) generated a feeling of familiarity in word learning tasks using conceptually related stimuli. They use Deese–Roediger–McDermott (DRM)-like stimuli (Roediger & McDermott, 1995), in which one critical non-presented item (e.g., SLEEP) is related to a series of other targets (e.g., BLANKET, REST, DOZE, DREAM, BED). At test, the non-presented item will feel familiar to the participant, and in an oft-replicated finding, a large proportion of people report a false memory for this item. Urquhart and O’Connor (2014) produced an analogue of déjà vu by adding a task at study whereby participants also had to report and note the number of words beginning with a certain letter, in our example, “S.” Because SLEEP was not presented, and no other word was presented that began with an “S,” it created a situation where the participants had knowledge about the study phase: “no words began with an ‘S.’” Thus, when presented the item SLEEP in the test phase, participants found the item very familiar but were also confronted with the information that they had not studied any words beginning with an “S,” and participants were likely to report having an experience of déjà vu, hence O’Connor and colleagues’ emphasis on the role of conflict in the déjà vu experience.

As a final note, déjà vu is a subjective experience where the key characteristic is its phenomenology and as such it is not only difficult to verify people’s experience, it is also possible that experimental inductions exaggerate the rate of déjà vu generation, or that participants confuse different types of phenomenological familiarity (O’Connor & Moulin, 2010). One concern is that participants report having had déjà vu even in a control condition. Central to this problem is the notion of demand characteristics, a response bias generated by social desirability (Orne, 1962), sometimes referred to as reactivity. The fact that participants are asked whether they have experienced déjà vu repeatedly adds to this possibility (e.g., Winkler, Kanouse, & Ware, 1982).

Jersakova et al. (2015) examined the possibility that reports of déjà vu were influenced by demand characteristics. They reasoned that participants could interpret being asked about whether they experience déjà vu as indicative that they *should* be experiencing it. This could arise with a participant who is unsure of what exactly she is experiencing or how to describe it. Also, it could be that the communication between the experimenter and participants about the exact nature of the studied experience is not clear. Cleary et al. (2009) report a cross-experiment comparison of déjà vu rates for otherwise identical experiments when they do and do not provide a definition of déjà vu. They observed changes in how the participants responded to their induction of familiarity, noting that “... participants’ pre-experimental notions about déjà vu may make them more inclined to equate déjà vu with standard feelings of familiarity” (p. 1087).

Jersakova et al. carried out online experiments with an ongoing continuous recognition task, reporting whether a given item was old or new. Interspersed with this task were questions about other experiences (notably déjà vu and the tip of the tongue [TOT]), which were manipulated between subject, and there was

also a control condition with no questions asked during the recognition phase. All participants had a common final post-experimental questionnaire that asked about the incidence of subjective experiences, the hypothesis being that asking frequently about déjà vu would lead to an elevated level of déjà vu reporting. Across experiments, between 33 and 58% of participants reported experiencing déjà vu or TOT – even if they had not been in a condition where they had been asked repeatedly about those conditions. Changing the definition of déjà vu or asking participants to bring to mind a real-life instance of déjà vu or TOT before completing the task had no impact on reporting rates in the post-experimental questionnaire. Jersakova et al. also compared their “laboratory” déjà vu and the real-world experience finding that participants rated salience, intensity, and emotionality higher for real-life déjà vu and TOT experiences as compared to their experimentally generated analogues.

Our aim here is not to question the outcome of experiments examining déjà vu formation, but rather explore how demand characteristics, and “pre-experimental notions” might influence responding of such a subjective and difficult to describe state. It should be noted that déjà vu research is not alone in possibly being influenced by demand characteristics – as we have noted here, for example, the expectancy of generating an involuntary memory also influences the rate of this experience too. However, all experimental approaches to déjà vu formation involve asking the participant directly if they have had déjà vu, and often on a trial-by-trial basis. This “provocation” of déjà vu is therefore somewhat at odds with our characterisation of déjà vu as spontaneous. Our point here is about the phenomenological spontaneity and unpredictability of the experience, which lends itself to the design used to provoke involuntary memories. Whilst we characterise déjà vu in the real world as being an unexpected, spontaneous experience, we do not yet know the extent to which it is experienced as surprising or spontaneous in experimental settings. It seems that two lines of experimentation might address this issue. The first would be to adapt the existing successful paradigms by Cleary and colleagues reviewed above such that questions are asked about the spontaneity and “automatic” nature of the experience. Second, we propose below an approach whereby a mundane ongoing task leaves space to experience spontaneous phenomena such as involuntary memories to complement the existing work.

2. Retrieval processes in involuntary autobiographical memory and déjà vu

Whereas IAMs are mostly discussed within the context of autobiographical memory (where recall tasks are predominant), déjà vu has been mostly discussed in the context of familiarity processes in recognition memory decision making. To reconcile these literatures, our hypothesis is that, given that both déjà vu and IAMs relate to our personal past, they reveal something about human retrieval processes, in general, and autobiographical memory retrieval, in particular. A complete account of retrieval should incorporate both these phenomena and demonstrate how they relate to each other. Importantly, while the idea that IAMs and déjà vu may be similar phenomena has been suggested in the literature (e.g., Bradley, Moulin, & Kvavilashvili, 2013; Conway & Loveday, 2015; Illman et al., 2012; Mazzoni & Hanczakowski, 2011), to our best knowledge, the present paper is the first that provides an account for the possible underlying mechanisms of both. Of note, in the only empirical work on

the relationship, Moulin et al. (2014) find that retrospective questionnaire reports of the frequency of déjà vu and IAMs in daily life are related, at least in older adults (mean age 71; $r[44] = 0.723$).⁴

2.1 Familiarity and feelings of fluency

The experimental induction of déjà vu converges on one theoretical entity: Familiarity. The implication is that retrieval is incomplete: Participants find a stimulus familiar but at the same time, are unable to retrieve its source (the Brown, Marsh, and Cleary experiments), or know that the familiarity is false (the Urquhart and O'Connor experiments). In terms of memory retrieval, as pointed out by Renoult, Irish, Moscovitch, and Rugg (2019), familiarity is often defined by default – it is what occurs in the absence of the rich evocative mental time travel which is experienced in episodic remembering (Renoult et al., 2019). For a definition of familiarity, we turn to the recognition memory literature, where it is described as a recognition memory decision-making process based on the evaluation of trace strength, a judgement of prior occurrence based simply on a feeling, rather than the retrieval of information (e.g., Yonelinas, 2002). A robust literature examining this trace strength account using signal detection models and receiver operating curve analysis exists (for a review, see Yonelinas & Parks, 2007; see also Delay & Wixted, 2021, for an example application to recognition memory making the case for a continuous process).

The focus here however is on phenomenology rather than recognition memory performance, and perhaps the idea of familiarity is most intuitive when we talk about first-person experiences, such as “the butcher on the bus”:

Consider seeing a man on a bus whom you are sure that you have seen before; you “know” him in that sense. Such a recognition is usually followed by a search process asking, in effect, Where could I know him from? Who is he? The search process generates likely contexts ... Eventually the search may end with the insight, That’s the butcher from the supermarket!

Mandler (1980, pp. 252–253)

In Figure 1, we propose that such phenomenological familiarity occurs when there is (temporarily) no content retrieved. It is felt in both involuntary and voluntary retrieval, but as we will explain below, it serves to guide and motivate retrieval processes: A sense of familiarity arises without any intention, but the subsequent search for corresponding information can be intentional. Before examining this metacognitive account in more detail, we quickly review the notion of familiarity in recognition memory.

Recognition of a cue in the environment arguably relies on two retrieval modes or processes, familiarity and recollection (e.g., Yonelinas, 2002). The neural and cognitive organisation of these entities is often debated (for an overview of this debate, see Moulin, Souchay, & Morris, 2013, and for a proposal for the integration of these ideas, see Bastin et al., 2019). Simons, Ritchey, and Fernyhough (2022) propose that early characterisations (such as Yonelinas’s [2002] account of a hippocampal mechanism and a simple threshold that gives rise to recollection) are helpful for explaining binary (laboratory-based) memory decision making. However, with more complex representations of the personal past, they propose that additional to this hippocampal reinstatement, cortical regions such as the angular gyrus are recruited to generate a precise representation of the episode (e.g., Richter, Cooper, Bays, & Simons, 2016). As such, they describe

remembering as drawing upon two different dimensions: Retrieval success and the precision of the details retrieved.

We emphasise separable *phenomenological* experiences of familiarity and recollection, and we are agnostic as to whether these map onto one or two underlying processes, although as will be seen, our ideas converge on the idea of a trace-strength account based on the continuity between familiarity and recollection. A nuanced view would be that complex interactions exist between networks traditionally thought responsible for familiarity (extra-hippocampal regions) and recollection (the hippocampus) such as proposed by Quamme, Yonelinas, and Norman (2007), with also the recruitment of an area around the left angular gyrus which yields the multisensorial information experienced in remembering as outlined by Simons et al. (2022). As proposed by Montaldi and Mayes (2010), we suggest that recollection and familiarity are related: "...each is a complex function, likely to depend on several different processes that are probably mediated by different structures that are functionally connected in a system" (p. 1294). Whereas the consideration of recollection and familiarity as *processes* favours an interaction between shared neural and cognitive processes in a continuum, for the experient, recollection and familiarity are separable signals which differ in their phenomenology and content.

The critical issue, as identified by Bastin et al. (2019), is that these signals are interpreted differently, and have consequences for retrieval. They propose an integrative model with a representational hierarchy and an attributional system. The first is a "core" system that stores specific representations based on interactions between a large network of computational operations from the whole brain, split into an entity representation system (guided by feelings of fluency – see below) and a relational representation system (responsible for binding and pattern completion) centring on the hippocampus. A second system acts on the current context and "translates content reactivation into a subjective experience" (p. 5) and includes input from networks in the parietal and prefrontal cortexes. Critically, their model rests on differences in the outputs and representations of their core systems, and an interaction between processes and computations, resonating with the Montaldi and Mayes "kinds" of memory proposal and our view about separable phenomenology. In autobiographical memory, the retrieval of specifics from the personal past is accompanied by a feeling of mental time travel or *autonoesis* (see Conway, 2009), directly comparable with the concept of recollection: A detailed representation of a moment in the personal past. The idea of the "self in past" (Conway, 2009; Klein, 2013; Marshall, Halligan, & Wade, 1995; Moulin & Souchay, 2014) is that it enables distinguishing between imagined, false, and actual events and gives episodic memory its particular character (Perrin & Rousset, 2014). Recollection gives the feeling of completeness, and the resolution of the feeling of familiarity. In the butcher on the bus, if we are able to identify the person we have seen, we resolve the feeling of familiarity and the feeling of recollection confirms successful retrieval. The phenomenology of remembering is indeed associated with high degrees of confidence (Dunn, 2004).

Recognition decisions made based on familiarity are proposed to be faster and more automatic than those based on recollection (e.g., Besson, Ceccaldi, Didic, & Barbeau, 2012; Hintzman, Caulton, & Levitin, 1998; Hintzman & Curran, 1994; Mandler, 2008; see Dewhurst, Holmes, Brandt, & Dean, 2006, for a discussion of this issue, and Berry, Shanks, Speekenbrink, & Henson, 2012, for a single-process explanation of this pattern of results).

In tasks where recollection and familiarity are compared by asking participants to make decisions based on the information they are able to retrieve, Yonelinas and Jacoby (1994) found that recollection decisions peaked at 800–1100 ms, whereas familiarity decisions peaked earlier at between 600 and 800 ms. In EEG paradigms (e.g., Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Rugg & Curran, 2007), two separate neural signatures are found for old/new decisions in episodic memory. The FN400 effect (between 300 and 500 ms, e.g., Curran, 1999; a negative potential in the frontal areas) is proposed to reflect familiarity processing, whereas the parietal old/new effect (sometimes called the late parietal effect) is a positive potential in the parietal area arising slightly later, after about 500 ms (e.g., Tsivilis et al., 2015).

Familiarity is often described as a low-level process working on perceptual inputs⁵ (Bastin et al., 2019). This part of the memory system tries to match, as quickly and effortlessly as possible, the contents of mental representations stored in memory with the current contents of perception; consistent with the classic neuro-anatomical view of the temporal-lobe memory system as being the last point in the ventral visual stream (e.g., Suzuki & Amaral, 1994). Viewed like this, familiarity decision making is the last stage in perception – once we have composed and identified a scene or environment, we can "read off" whether we have encountered it before. More recent neuroanatomical accounts of memory retrieval propose a similar organisation. Moscovitch (2008) describes two phases in episodic retrieval. A first stage involving the hippocampus is non-conscious, based on a rapid interaction between a retrieval cue and a stored representation, whereas a second stage, featuring a more complex interaction between the hippocampus, prefrontal, and parietal cortex, is required for the retrieved information to be re-experienced. These accounts are inspired by the notion of an organisational hierarchy of connectivity (e.g., Margulies et al., 2016). Irish and Vatansever (2020, p. 47) propose a "...concrete to abstract representation spectrum along modality-specific and default mode brain networks which collectively support rich and detailed memories," such that rather than specifying different stores or types of information as in the standard episodic/semantic distinction (see also Renoult et al., 2019), the memory system emerges as part of a larger scale network with retrieval of specific events from the personal past as being a type of interrogation of the stored representations based on a common network, rather than something restricted to a specific zone, region or process in the brain. As such, memories might be thought of as reflecting a dynamic pattern of local to regional connectivities, varying along the continuum from abstract to specific and more or less involving the sensorimotor regions (Irish & Vatansever, 2020). For an up-to-date review of the neural processes in the subjective experience of remembering taking on board both the large-scale networks and traditional "regions" involved in memory, see Simons et al. (2022).

The existing cognitive frameworks resonate with the neuroscientific view, but discuss phenomenology in a way which is somewhat lacking in neuroscientific accounts (see Renoult et al., 2019). In Whittlesea's SCAPE (Selective Construction and Preservation of Experience) model (Whittlesea & Williams, 2001), the memory system is continuously trying to make sense of its inputs, so that it can interpret any signals arising from low-level processing of the environment. Familiarity is thus a subjective feeling arising from the fluent processing of a stimulus. This theory draws upon various illusions of familiarity: Where the fluency of processing leads to the (erroneous) decision that we have seen something before, or to "know" an answer (e.g., Goldinger & Hansen,

2005; Newman, Garry, Bernstein, Kantner, & Lindsay, 2012; Whittlesea & Williams, 1998). Phonological and structural regularity can be used to induce feelings of fluency which are misinterpreted as meaning we have seen the information before. For example, in the *hension* effect, people are proposed to be more likely to make false positives for pseudowords such as *hension* (which is regular and resembles a real word) than *stofwus* (for a critique, see Cleary, Morris, & Langley, 2007). The idea is that the ease at which they read the word is misinterpreted as it being an item from a study list. Golding and Hansen push the idea one step further: Participants respond to meaningless subliminal vibration cues, interpreting them as a signal they have seen the word before. There is some debate about whether the critical element is “surprise” in false memories, and it seems unlikely that Whittlesea’s view could explain all kinds of false memory. For instance, in the DRM paradigm, the critical lures for which participants have false memories are *expected* as targets, and so the false memory cannot be described as surprising (Karpicke, McCabe, & Roediger, 2008). Moreover, Cleary et al. (2007) point out that structural regularity usually leads to benefits in recognition memory, not an increase in false positives, suggesting that empirical support for Whittlesea’s ideas on false memories is weak. Cleary et al. (2007) found that the *hension* effect disappeared completely once controlling for inter-item similarity in the recognition test materials, and so interestingly, this effect which formed the basis of the original theorising receives little empirical support. The role of phenomenological fluency in recognition memory is something which will need continued evaluation, even though the basic idea, that the ease of processing leads to an attribution of prior exposure, is well established (e.g., Jacoby & Dallas, 1981; but see also Berry et al., 2012).

Whilst there is not unequivocal empirical support from recognition memory studies, Whittlesea describes an interaction of attributions of memory and processing fluency, something that is critical for thinking about how cognitive “feelings” guide retrieval processing (Bastin et al., 2019; Moulin & Souchay, 2014). The idea that we adopt here is that the current context generates a top-down expectation of processing fluency. When an expectation of fluency is violated, it triggers a search for as to why that has arisen. This is the search triggered in the case of the butcher on the bus; the person feels intensely familiar precisely because the source of the familiarity is unknown. In this way, we align our reasoning with Whittlesea’s observation that strangers feel familiar, but our friends do not (Whittlesea & Williams, 1998): The phenomenology of familiarity is produced by a mismatch between processing fluency and the expectations about fluency. The spontaneity and phenomenology of the experience are one and the same: It feels familiar because it was unexpected.

Fluency is also a variable of interest in IAM studies. Sanson et al. (2020; see also Barzykowski et al., 2021b) investigated the influence of fluency on classification of intention in a paired-associates design. Participants studied a series of phrases, half of which were accompanied by a phrase-related photo. During retrieval, participants were presented with the phrase only and then instructed to retrieve additional information about it. Crucially (Experiment 3; Sanson et al., 2020, p. 12), participants reported if they experienced retrieval as spontaneous (“the information came to mind without intent”) or deliberate (“the information was brought to mind intentionally”). Although they were explicitly instructed to intentionally retrieve memories, fluency overruled intentionality when classifying

their memories. When participants intentionally, yet fluently, brought a memory to mind, they were prone to judge the retrieval as spontaneous. Similar results were obtained in study by Barzykowski et al. (2021b) who investigated the effect of retrieval effort while classifying memories. They created experimental conditions that maximise the probability of one type of retrieval (e.g., involuntary or voluntary) while minimising the probability of the other. Participants recalled autobiographical memories in each condition, retrieval time was measured, and they classified their memories as either voluntary or involuntary and rated them on retrieval effort. There were four categories of memories: Experimentally defined voluntary and involuntary memories with an objective measure of effort (retrieval time) and subjectively classified involuntary and voluntary memories with a subjective measure of effort. Similar to Sanson et al., the majority of memories were classified as involuntary, whether they were retrieved in the experimentally defined voluntary or involuntary condition. Interestingly, the relationship between the rating of subjective fluency and objective retrieval time was only significant in the voluntary conditions, indicating that people are more aware of subjective effort when deliberately trying to retrieve a memory than for an involuntary memory. In contrast with existing theories, these results indicate that the subjective feeling of fluency is more important than intention for the involuntary–voluntary distinction. Such experiments suggest that low-level feelings are used when making explicit decisions about memory.

These findings question the interpretability of existing studies on involuntary memories in which fluency has not been controlled for. Sanson et al. (2020, p. 2) argued “[...] subjects themselves classify their memories as voluntary or involuntary, and so the validity of what scientists know about involuntary memories rests upon the validity of this task. How well do subjects perform this task, then? Here, we show that, at least under certain circumstances, the answer is not very well.” Our view is more moderate: Classifications are internally consistent in terms of retrieval latency and retrieval effort, especially during voluntary retrieval. At the same time, it is much less of an issue during involuntary retrieval, since the vast majority of participants’ classifications align with the classification set by the experimental condition. As such, these findings do not undermine previous results on involuntary and voluntary memories but rather extend our understanding of retrieval. Accordingly, the involuntary–voluntary distinction is not as clear-cut as the conceptual foundation of intentional versus non-intentional retrieval (or indeed retrieval with or without cognitive control) suggests. As such, there are no involuntary/voluntary memories but a memory representation that is constituted by certain factors operating during retrieval, an attributional process. We see retrieval as a dynamic process consisting of interactions between different dimensions (e.g., intention, retrieval effort, memory accessibility, and monitoring processes) that all together make up the particular “flavour” of a retrieved memory.

We see familiarity as metacognitive; a feeling that is generated from the fluent processing of a stimulus.⁶ It is a signal which guides memory processing – orientating us to information in the environment which we have already encountered. As such, it can be described as an epistemic feeling (Moulin & Souchay, 2014). Arango-Muñoz (2011) defines epistemic feelings as quick-acting intuitive processes, based on how things feel, and déjà vu is often used by philosophers as an example of an epistemic feeling (Arango-Muñoz, 2014; Bortolotti, 2010; de Sousa, 2009; Gerrans, 2014). Epistemic feelings neatly map onto a set of metacognitive phenomena and paradigms that are better researched and

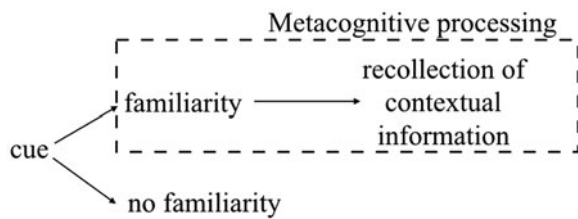


Figure 2. Representation of the metacognitive evaluation of familiarity and recollection in recognition memory decision making.

understood than déjà vu, such as the TOT state (Greely, 2021) and the feeling of knowing (FOK; Souchay, Moulin, Clarys, Taconnat, & Isingrini, 2007). These are instances of metacognitive awareness at the moment of retrieval failure (Fig. 1). When familiarity is high and retrieval is incomplete, this will lead to such feelings, even for the personal past. Critically, in the TOT (or the FOK) for the personal past, the experient must evaluate retrieval as plausible, since otherwise the experience would be one of conflict, possibly resulting in déjà vu.

Souchay and Moulin (2009) synthesised the existing proposals into a schematic representation about the relationship between familiarity and recollection (Fig. 2). Various similar proposals exist in the literature (see Hintzman & Curran, 1994; Koriat & Levy-Sadot, 2001; Reder, 1987; Reder & Ritter, 1992).⁷ It explains the search for information given a cue in the environment, and pertains most clearly to the FOK. In the FOK, a rapid initial feeling of familiarity with a cue triggers the attempted retrieval of episodic detail about the cue, and any associated information. In the absence of any contextual information, people can make accurate predictions of future recognition based on the strength of familiarity for the cue, and what “partial information” about the cue or searched-for-target comes to mind. Such research shows that even when information is not currently retrievable, outputs from the recognition memory system (in particular fluency) can be acted on metacognitively in order to accurately predict the likelihood of subsequent retrieval, suggesting the existence of a sequential arrangement as shown in Figure 2, which is why it is helpful to think of familiarity and recollection as lying along a continuum.

In the context of IAMs, it is possible to retrieve information directly without feeling anything until the moment there is a fully formed mental representation. We would argue that this is due to there being very little delay between the first familiarity stage and the complete retrieval. Memory retrieval in this case feels spontaneous because the subject had both the feeling of familiarity and recollected specifics at approximately the same time. Familiarity is most strongly felt when we are not able to retrieve contextual specifics. In the butcher on the bus, familiarity is high but the recollected specifics are (momentarily) absent, whereas in déjà vu, the familiarity is similarly high, but a metacognitive evaluation favours the interpretation that the familiarity is false or impossible. Here we suggest for the first time that this is based on access to information about the personal past.⁸

2.2 Autobiographical retrieval mechanisms

Here, we examine in more detail autobiographical memory retrieval processes with a focus on metacognition, and propose a threshold mechanism. According to Conway and Pleydell-Pearce (2000; for later modifications, see also Conway,

2008, 2009; Conway & Jobson, 2012), autobiographical memory consists of a hierarchical network of interconnected nodes that differ in terms of their level of specificity. At the bottom are fragments of events with specific sensory details (e.g., details experienced when riding a bike for the first time). Such vivid information is further connected to super-ordinate levels of general events (e.g., *riding a bike to work*), common themes (*riding, commuting to work*), and, at the highest level, important periods in one’s life (e.g., *when I moved to France*). The activation of autobiographical information spreads across the network resulting in the construction of a particular memory. Such activation may be elicited by different types of cues, internal or external. While voluntary autobiographical memories are the result of a top-down search process that eventually arrives at an episode, directly retrieved (i.e., involuntary) memories are thought to circumvent the search process and enter consciousness very quickly. Conway and Pleydell-Pearce argued that fragments of memory representations are constantly activated at the bottom level of the hierarchy, but the vast majority of such memories never reach consciousness due to being suppressed by constantly operating executive control system. Moreover, some low level and/or mundane information will be completely in line with expectancies (cf. Whittlesea’s SCAPE) and will not be raised to awareness. Only some of these activated memories, especially those that are consistent with current self-goals, may occasionally get through this executive control mechanism and reach awareness.

This implies that IAMs occur as a result of two concurrent processes, namely spreading activation and inhibition that works against such activation (see Ball & Hennessey, 2009). The retrieval of IAMs depends on absent or inefficient inhibitory control. For instance, when entering a hotel room for the very first time (i.e., we have never been to the city, in general, nor to this hotel chain), the environment is being continuously processed and automatically matched with any corresponding representation in memory. Aside of deliberative searches in memory cued by conscious thoughts (e.g., being asked by the receptionist if we have ever stayed at the hotel before), we propose that there are several possible outcomes, which lie on a continuum: (a) An involuntary memory may be triggered, (b) one may get only a general feeling of familiarity, or (c) the room does not trigger any retrieval process. The key premises delivered from this model are that (1) autobiographical memory information is constantly and automatically activated by cues and that (2) only some of them may enter awareness.

Given that IAMs occur automatically in response to incidental external and internal cues, one may ask why are we not constantly flooded by memories? According to the principle of cue overload, it is most likely that a cue will match several past events. Berntsen (2009) proposed a mechanism of cue-item discriminability, defined as “how easily a given cue isolates an item” (Rubin, 1995, p. 151 as cited in Berntsen, 2009, p. 107; see also Norman & Bobrow, 1979). That is, the more events that are associated with a particular cue, the less efficient this cue will be in triggering any one of them. Berntsen, Staugaard, and Sørensen (2013) examined the relationship between cue and memory and were able to predict the occurrence of an involuntary episodic memory based on a manipulation at encoding. Across four experiments, they found that involuntary episodic memories were retrieved more often in response to unique compared to repeated cues, in keeping with the principle of cue-overload.⁹

The déjà vu field has yet to manipulate such variables, although a neuroscientific account relating déjà vu to novelty

mechanisms has been advanced by O'Connor, Lever, and Moulin (2010). At least in clinical samples, delusional beliefs of prior occurrence are more likely to appear in novel locations and events (i.e., environments with unique cues) than those that are familiar. Our prediction is that déjà vu would arise in situations where the cues are relatively unique, since in situations where there is cue overload, the attribution will be that any sense of familiarity is normal given the current situation.

It can be argued that cues that are unique and distinctive are infrequent in daily life and most cues map onto many different memories and events. Given how frequent IAMs are, it must be the case that other processes are at play. In response, it has been suggested that priming processes (i.e., increasing the activation of a memory information by prior encounter with the contents of memory representation) *may* not only increase cue-item discriminability allowing for efficient activation of a particular involuntary memory (e.g., Barzykowski & Niedźwieńska, 2018b; Mace, 2005) but also increase the familiarity of a given cue.¹⁰ It is thus possible that on some occasions, a new event/setting may consist of cues that map onto many different past memories, or even that there is a relevant configural, contextual similarity between the current situation and some past events (exactly as proposed in the Gestalt Similarity hypothesis; Cleary et al., 2009). While these cues/contexts are not efficient in triggering any given memory, it is nonetheless possible that some of them may induce the feeling of familiarity, leading to the experience of déjà vu. This would be the case if the memory is vague, or non-specific: Associated with many previous events/information in the AM organisational hierarchy and unable to enter awareness. The activation is not strong enough to trigger a given memory, but it is strong enough to trigger a feeling of familiarity. This suggests that while on some occasions the memory content (e.g., *being in a similar hotel room in Rome a few years ago*) does not enter consciousness, leaving us without a straightforward reference to our past, the feeling of familiarity still appears. Conway and Pleydell-Pearce (2000) assumed an inhibitory control mechanism suppresses irrelevant memory information, preventing it from reaching consciousness. Findings from several recent studies (e.g., Barzykowski et al., 2019b, 2021a, 2022) do not provide full support for the existence of such a direct control mechanism,¹¹ and favour an account based on a retrieval threshold (see Barzykowski & Staugaard, 2016, 2018; Barzykowski et al., 2019a).

In recognition memory decision making, activation/threshold models have long been the norm, stemming from signal detection theory (e.g., Banks, 1970; Wixted, 2007). In autobiographical memory, it has recently been proposed that in order to reach consciousness, a memory must pass an awareness threshold that determines the minimum amount of activation for awareness. Entering awareness can be achieved either when a memory reaches levels of activation that is greater than a given threshold, or when the threshold is lowered by some additional factors. For activation, for instance, it may be easier for a phenomenologically pronounced and intense memories to pass the threshold because they may be especially good at drawing one's attention (see also Barzykowski & Staugaard, 2016, 2018; Barzykowski et al., 2019a). For the threshold level, it is argued that it may be momentarily modified by different factors (e.g., expecting memory to happen, placing the focus of attention on one's stream of thoughts), which would consequently modify the experience of otherwise less accessible memories. The threshold hypothesis states that, while both highly and poorly accessible memories can be retrieved either voluntarily or involuntarily, the processes

operating during memory retrieval can influence the frequency of each type of retrieval by modifying the awareness threshold. Barzykowski and Staugaard (2018; also, Barzykowski et al., 2019a) proposed these processes to be intention (i.e., wanting to retrieve a memory) and selective monitoring (i.e., expecting a memory to appear). This idea has been verified in several empirical studies (e.g., Barzykowski & Mazzoni, 2022; Barzykowski, Skopicz-Radkiewicz, Kabut, Staugaard, & Mazzoni, 2023; Barzykowski & Staugaard, 2016, 2018; Barzykowski et al., 2019a, 2021b). For instance, involuntary memories compared to voluntary memories were rated by participants as more phenomenologically accessible; namely, more clear, unusual, personally relevant, important, recent, and more frequently rehearsed in the past (e.g., Barzykowski & Staugaard, 2016; Barzykowski et al., 2019a). These findings corroborate the idea that when not being actively engaged in memory retrieval, a memory has to be phenomenologically sound to pass the threshold. Also, Barzykowski and Niedźwieńska (2016) demonstrated that involuntary memories retrieved in a selective monitoring condition (i.e., when one is expecting a memory to appear) were less clear, less detailed, less vivid, and elicited less intense physical reactions compared to memories retrieved without selective monitoring. In another study, selective monitoring decreased the personal importance and emotional intensity of memories (Barzykowski & Staugaard, 2018). Finally, these results pertain not only to the phenomenological characteristics of memories but also to their objectively observable descriptions (Barzykowski et al., 2023). When looking at the descriptions of memories, *on average* involuntary memories were indeed more accessible (i.e., rated by independent judges as having more indicators of memory accessibility such as being scored higher on the emotional impact scale, etc.) than voluntary memories and intentionality lowered the awareness threshold and increased the number of low-accessibility memories.

Note that these threshold processes are metacognitive, about expectations and intentions. Our proposal is that in complete retrieval, there is little need to reflect too much on the process of retrieval: Information comes to mind, and is recollectively experienced, whether or not it was sought for. Naturally, the recall of contextual specifics may even justify to the experient *why* the information has come to mind, for example, "it was in this post office where I got my foreign currency last year." However, in cases where retrieval is incomplete or low-level signals violate expectations (i.e., finding something familiar in a novel location), then intense feelings are generated.

Autobiographical retrieval is therefore determined by a complex interplay between factors. Intention and monitoring can be conceived as processes that enable access to otherwise less accessible memories. The most important assumption of the threshold hypothesis is that because of certain phenomenological properties (e.g., emotional intensity, clarity, vividness, personal relevance, unusualness), some mental contents may be especially good at drawing one's memory-related attention and, thus, they may pass the awareness threshold more easily and be more likely to be reported. However, on occasions where we may not be aware of a given memory content, but where activation is sufficient, we can still feel familiarity. If the familiarity is logically impossible and unexpected, it will trigger a search in memory to explain the feeling. Thus, both déjà vu and IAMs start from the same point in shared retrieval processes.

We turn our attention to how these processes can be split down into their component parts. Autobiographical memory

retrieval has the following four stages (e.g., Barzykowski & Mazzoni, 2022; Wilckens, Erickson, & Wheeler, 2012): Pre-retrieval, retrieval, post-retrieval, and retrieval outcome report stages.¹² The pre-retrieval stage is associated with any cognitive processes that either facilitate or impair retrieval. For example, during this stage, an individual may be in “retrieval mode” in which “the cognitive system is prepared for or expects memory construction and recollection” (Conway, 2001, p. 1379). The effect of priming would occur here, enhancing the likelihood that the memory will be triggered and/or will enter a person’s awareness. The retrieval stage relates to the forming and developing of an autobiographical memory, but without explicit self-reflection; namely, a given memory might have been formed but one may not be explicitly aware of it yet. During this stage, a memory is triggered by and/or accessed via a given cue. Once the memory is formed, during the post-retrieval phase, people may become aware of having the memory in mind. Thus, this stage relates to the ability to, for example, extract autobiographical content from the stream of consciousness to explicitly become aware of having a memory that is autobiographical. One is fully aware that an autobiographical memory was actually retrieved, and in voluntary retrieval, if the memory content meets the given criteria, then the search may be terminated. In the last stage, the retrieved memory may be shared with others and reported by giving a verbal account of the content.

This four-stage account highlights two main factors, the first being intentionality during the pre-retrieval phase, the second being a metacognitive process (see Koriat & Goldsmith, 1996; Mazzoni & Kirsch, 2002). Koriat and Goldsmith, inspired by signal detection models, describe a situation whereby the subject can decide whether to report retrieved information (p. 181): “people monitor the subjective likelihood that an item of information that comes to mind is correct and then apply a control threshold to the monitoring output in order to decide whether to volunteer that item or not.” Their emphasis, necessitated by the use of a forced/free report recall procedure (see Goldsmith & Koriat, 1999), is on the participant evaluating the quality of information before choosing to output it, whereas our proposal is that internally such a process also exists to motivate and regulate the search for information for the subject.

When looking at these retrieval stages, we argue that both déjà vu and IAMs are natural forms of involuntary and spontaneous processing of cues that can be found in one’s surroundings, and that memory information when retrieved is not only accompanied by the exact memory content but also the feeling of familiarity that, importantly, may arise and/or be experienced separately from each other. In particular, déjà vu can be thought of as arising at the retrieval stage, and before the person is aware of having any content in mind.

A critical issue is to understand why in some cases, retrieval is incomplete and déjà vu arises rather than the retrieval of an IAM. Here, we advance two hypotheses which are not mutually exclusive and may map onto the top-down/bottom-up descriptions of déjà vu in the literature. These are proposals based on cue familiarity/trace strength and the metacognitive evaluation of cues in the environment. It is beyond the scope of this paper to delve too deeply into retrieval failure per se on a mechanistic level, but it is important to note that the dynamic patterns of inhibition and activation seen in episodic memory (e.g., Anderson, Bjork, & Bjork, 1994; for a review, see Kuhl & Wagner, 2009) are likely to contribute to the failed retrieval referred to here. For instance, it is known that competition between similar cues leads to an

inhibition process whereby the retrieval of one memory actively inhibits the retrieval of another. It is not impossible that information is not simply forgotten or lost according to this process, but that it gives rise to phenomenological states such as referred to here, something that would warrant further investigation.

Our first hypothesis is that there is only weak activation of the associated cues and memories from the environment, and thus there is not sufficient activation for any content to enter into consciousness. This is not unlike the experience of information in the TOT (developed in detail below), where knowledge or lexical information is partially activated and frustratingly will not come to mind, even if we are aware of its existence. In this weak version of déjà vu, a parallel is seen in autobiographical memory. A cue reminds us of something stored in autobiographical memory, but the cue is undetected and does not yield sufficient activation, giving a feeling of familiarity without a known source. Accordingly, it would be expected that déjà vu will most likely occur in a context that is rather familiar/known, in contradiction to what we proposed above when discussing cue discrimination. Whilst Brown and Marsh (2008) compared mundane and distinctive scenes in their experimental analogue of déjà vu, finding that mundane scenes were more likely to generate feelings of having visited a place before, this rather common-sense interpretation of déjà vu in mundane, familiar locations has yet to be tested (although questionnaire studies suggest that in healthy groups, déjà vu happens more in familiar than novel contexts; see Moulin, 2018). One possibility to test this idea in experimental analogues of déjà vu is to further probe the déjà vu once experienced. For instance, one may amplify the weak activation of the associated cues and memories by lowering the awareness threshold and/or drawing the attention to the possible sources of déjà vu. This way, the feeling of familiarity caused by weak activation of memory content could be overcome leading eventually to memory content retrieval; that is, reinstating cues experimentally may shift the experience from déjà vu to a memory of the past. This weak activation account also suggests that the déjà vu experience may be more likely to arise when one is engaged in attention demanding activities where ongoing task performance elevates the awareness threshold making weakly activated memories difficult to enter consciousness.

Our second hypothesis centres on metacognitive evaluations and posits that déjà vu is felt precisely because it is unexpected, aligning with the top-down view. In certain contexts, the expectation of familiarity should be low. In such cases, any involuntary process of retrieval, especially when it does not result in any retrieved information will be unexpected and will lead to a metacognitive evaluation which will be raised to consciousness. Until this feeling of familiarity is resolved or explained, it will grow and the experiment may well search for a reason “why” they have this feeling. Experimentally, it would be relatively simple to manipulate expectations about the likelihood of finding information familiar. In Cleary-type recognition without identification experiments, the idea that expectations about familiarity lead to the feeling of conflict should mean that déjà vu is more intense and more frequent when the expectation is lower. This could either be manipulated by giving explicit instructions to this effect to participants or manipulating the amount of similarity between the initial scene and its analogue in the recognition phase. It would be intense when the context is surprising, such as in the unrelated experiments conceit (e.g., Lakin & Chartrand, 2003): A first study phase is conducted in the context of one task and then the test phase is presented as a separate experiment. This

design could assess the role of the experimental context (either presented as the same experiment or a separate experiment) on the report of déjà vu.

There is little existing research to discriminate between our two hypotheses, but Urquhart and colleagues successfully manipulated both the novelty and the activation (which they described as “familiarity”) of the critical non-presented stimulus. For novelty, they manipulated how specific the letter string was in referring to the critical lure (e.g., monitor for words that begin “SLE” in the list, vs. monitor words that begin with “B”). In the test phase, the string SLE was highly novel because it did not correspond to any items presented in the study phase, whereas the less specific “B” string applied to two words in the studied list, *blanket* and *bed*. For activation of the cue, they used results from previous versions of the DRM task: Some lists of associates are more likely to activate the critical lure than others: Thus they had more strongly and more weakly activated critical lures. In their experiment, the recognition memory results replicated the expected patterns. However, for the rates of experiencing déjà vu, there were no effects or interactions involving activation: The strength activation of the critical lure in the presented lists had no bearing on the generation of déjà vu. Thus, in this case, a weakly activated associate was not more likely to generate déjà vu than a more strongly activated one. The effect of novelty was however significant: Participants were more likely to experience déjà vu when the critical lure was unambiguously impossible to have been viewed in the study phase due to its unique string.

Our explanation of déjà vu as being “unexpected” would involve processes downstream, corresponding with the post-retrieval stage. During this stage, one becomes aware of having either an involuntary memory in mind or having a feeling of familiarity. More precisely, during both the post-retrieval and final stage, the unexpected feeling of familiarity without having any particular memory in mind may start metacognitive processing (e.g., thinking of plausibility of the event, etc.) leading people to explicitly experience déjà vu, especially in cases when the current situation is judged as novel and, crucially, “encountering it in the past is in fact highly implausible” (Mazzoni & Hanczakowski, 2011, p. 104). This would be the conflict described by Urquhart and O’Connor (2014): Familiarity is generated bottom-up from the environment, but higher-level autobiographical belief or knowledge opposes the feeling. This should be a rare occurrence, since as Conway (2005) proposes, autobiographical remembering is typically coherent with current goals and processing, but this would be a situation where the experience is incoherent. An interesting way to test this idea would be to re-evaluate reported déjà vu experiences in questionnaires according to self-reports of implausibility and novelty, on one hand, and also their stability across time. More precisely, although a given situation may provoke déjà vu, it may do so because in fact, there is not sufficient access to information relating directly to a past situation. However, such access may be restored later, and this additional information may change the evaluation of what was initially experienced as déjà vu, raising the question of the stability of déjà vu over time. In the only work related to this idea, Milton, Butler, and Zeman (2011) report a case of epileptic amnesia where the spontaneous retrieval of previously inaccessible autobiographical memories was “heralded” by a period of déjà vu. Finally, we may also hypothesise that if déjà vu experiences arise in response to the unexpected feeling of familiarity without having any particular memory in mind, then increasing the expectancy of a feeling

of familiarity in response to the current situation should lower the frequency of reporting déjà vu experiences.

In our proposal, memory retrieval is not an all-or-nothing process, and numerous forms of partial or incomplete memory retrieval exist.¹³ Notably, memory content and a feeling of familiarity are separable (e.g., Ryals & Cleary, 2012). Studies on the TOT phenomenon (e.g., Brown, 1991; Heine, Ober, & Shenaut, 1999; Schwartz, 2001, 2002, 2011) demonstrate that on some occasions while we cannot access memory content (e.g., *trying to recall the last time we were in cinema*), we have a strong feeling of familiarity, with the attribution that this memory exists somewhere. Whilst this phenomenon is usually discussed in terms of linguistic or semantic processes, our proposal is that it occurs for the personal past too. Thus, it may occur for a given cue (e.g., a colourful beach *ball*): There is a strong feeling of that there is something in our past relating strongly to this cue or that this cue reminds us of something that cannot be yet recalled.

In Figure 1, we propose a flow chart summarising the possible memory phenomena on three dimensions: (1) Retrieval intentionality (involuntary vs. voluntary), (2) memory content accessibility (the presence vs. absence of memory content), and (3) phenomenological familiarity (feeling familiar or not). It shows that TOTs/FOKs are the result of unrecalled content despite having strong feelings of familiarity. They can arise both when retrieval is intentional and unintentional, and in this way, they act as a bridge between more controlled and less-controlled processes. That is, an involuntary retrieval may provoke a sense of familiarity, that once detected will lead to an intentional retrieval or at least a wilful memory search. In this way, metacognition plays a functional role in recruiting intentional retrieval processes to something for which retrieval was cued in an involuntary manner.

Figure 1 makes the case that familiarity is something that is experienced (or not) on the way to retrieving content, and that there is something of a sequential relationship between feelings of familiarity and the retrieval of content. We think that this is necessarily the case in failed retrieval, where there is a phenomenological response despite there being a lack of content, but this is not necessarily the case in fluent, successful retrieval where the presentation of content to consciousness is so rapid as to feel direct. The organisation may be sequential, but it is also recursive: Feelings of familiarity and content retrieval may jump back-and-forth in a manner inspired by the large-scale synchronisations of activation in the DMN; more precisely, this phenomenology-content loop is inspired by the interactions between retrieval and control networks stated in various models of memory retrieval described above (e.g., Eichenbaum, 2017; Irish & Vatansever, 2020; Sekeres et al., 2018).

Our view is that one underlying retrieval process yields two different experiences: Phenomenological familiarity is experienced when retrieval is incomplete, but it is not experienced when retrieval is fluent and routine. That is why our friends do not feel familiar, since we are able to rapidly retrieve specific information about them. That is, high trace strength memories can be accessed directly without giving rise to phenomenological fluency. In Figure 1, voluntary memories can be retrieved with familiarity or more directly, but involuntary memories never give rise to familiarity during retrieval. The study of déjà vu and IAMs helps us arrive at this conclusion since we need to understand why there is sometimes unbidden retrieval of the personal past without giving rise to any particular phenomenology, and on the other hand, in déjà vu, there is the opposite: An

intense feeling of familiarity with the conclusion that the familiarity is erroneous, and devoid of content.

In sum, both déjà vu and IAMs are products of involuntary retrieval but, as presented in Figure 1, while IAMs are accompanied by memory content, déjà vu lacks the access to memory content and the feeling of familiarity is judged as implausible (if judged plausible, then one would most likely experience FOK/TOT).¹⁴ An interesting consequence emerging from Figure 1 is the possibility of differentiating involuntary and voluntary instances of FOK/TOT, which for now in our characterisation are the same experience whether or not they are encountered in intentional or unintentional retrieval. Although they have not been explicitly contrasted with each other in previous studies, such a distinction seems to us to be possible. Finally, all these phenomena (involuntary/voluntary memories, déjà vu, FOK/TOT) can be described along the same retrieval dimensions and so the integration of déjà vu into autobiographical retrieval has opened up new way of thinking about and classifying the range of possible retrieval experiences. We discuss this now, pointing out where research is needed.

3. What is already known and what still needs to be known

Whereas déjà vu research has been aligned with familiarity processing, IAMs have been studied from the viewpoint of autobiographical recall. It seems to us that to reconcile many naturalistic phenomena, a more circumspect view is needed. For instance, in déjà vu, the role of conflict favours the view that a feeling of retrieval is at odds with what is known or *recalled* about the current situation. Very little in cognitive science speaks to our rapid ability to reject things as unknown or unexperienced (but see Kolers and Palef [1976], “knowing not”). Déjà vu cannot simply be only about familiarity but also about the failure of the familiarity to reproduce any stored representation which can explain it: Recall (or at least attempted recall) is involved in déjà vu. Likewise, a stereotypical view of involuntary memory is that information arrives in consciousness unbidden, and it is as such uniquely a recall phenomenon. However, as we have seen, it is likely that cue activation and information in the environment drives this recall, however effortlessly, and the view we present here is that there should be a continuum between finding something familiar in the environment and spontaneously recalling information related to it. Along this continuum are varying degrees of our three dimensions: Intentionality, accessibility of content, and the feeling of familiarity. The classification by the experient of a retrieval as being déjà vu or an IAM is a metacognitive evaluation applied at a later stage to more-or-less usual memory processes, which like low-level familiarity mechanisms, we consider to be somewhat permanently active, scanning the environment for novelty and familiarity (in line with Bastin et al.’s [2019] proposition). This, of course, implies a modulation between responding to the external environment and generating internal thoughts; a fundamental property of human cognition. In cognitive psychology, this is captured in the difference between internal and external attention (Chun, Golomb, & Turk-Browne, 2011), and in cognitive neuroscience, this issue is represented in the discussion of the role of the DMN in “perceptually decoupled thought” (Smallwood et al., 2013; see also Kvavilashvili et al., 2020).

One possibility is that purely spontaneous retrieval without feelings of familiarity and metacognitive evaluations of intentionality should be related to greater activity in the DMN, such as

medial pre-frontal cortex and the posterior cingulate, whereas stimulus-driven retrieval accompanied by familiarity, even when the retrieval of content is low, should be indexed by higher parahippocampal/medial temporal-lobe activation. Of note in Urquhart et al.’s (2021) fMRI investigation of conflict-based déjà vu, undetected false memory decisions (false positives to unstudied critical lures) reveal activations which are indistinguishable from genuine memory decisions (see also McDermott, Gilmore, Nelson, Watson, & Ojemann, 2017). It remains to be seen whether the prefrontal activations seen for decisions where conflict is detected are due to modulations of internal/external attention or align more broadly with metacognitive processes which too are proposed to lead to activations in prefrontal cortex (e.g., Morales, Lau, & Fleming, 2018).

Secondly, although both IAMs and déjà vu experiences have been studied for several years, one basic but yet central aspect of these two spontaneous phenomena was surprisingly overlooked; namely, their phenomenology. Therefore, there is still a need for experimentally oriented systematic studies to further examine exactly what it is *like* to have IAMs and déjà vu. Are they experienced similarly to each other and only the presence or lack of the content makes them distinct or, alternatively, are there more differences in the way they are phenomenologically experienced? It would be also interesting to examine how these two experiences are described by people. This leads to an intriguing question of the role of metacognitive beliefs in experiencing these spontaneous instances (for similar studies on mind wandering, see Zedelius, Protzko, & Schooler, 2021).

In addition, there is a need to examine the role of conflict in déjà vu. This “fact checking” control process is the thing that is hardest to reconcile with existing involuntary memory theory. We are obliged to propose that a later, metacognitive stage is able to reject cues in the environment as being false. In our schematic (Fig. 1), chief among the higher-order evaluations in memory retrieval is plausibility. In implausible situations, we are able to quickly reject the current situation as being novel or unique, when this happens quickly and decisively, we imagine that it is what generates déjà vu, as opposed to a search in memory for a prior event. Given the notion of the continuity between imagination and memory retrieval, in counterfactual memory (e.g., De Brigard, Spreng, Mitchell, & Schacter, 2015) or in representations of the future (e.g., Schacter et al., 2012), it seems important to identify the control mechanisms that allow us to experience memories as veridical or as having occurred in the past. Are these the same control mechanisms implied in the conflict generated in déjà vu? It has been proposed that déjà vu acts to correct faulty or overactive feelings of familiarity, and it follows that if one is less likely to experience déjà vu, then one should be more likely to falsely recognise situations and events.

Our proposal rests on the ability to judge the plausibility of retrieved material and even the plausibility (or likelihood) that the present situation should generate memories or feelings of familiarity. In comparison to investigations of familiarity or recollection, this is an underdeveloped field, and in general, there is a lack of studies evaluating metacognition in autobiographical memory (although plausibility evaluations have gained some interest, for example, Scoboria, Mazzoni, Kirsch, & Relyea, 2004). One way of rapidly judging plausibility online during retrieval would be by conceptual stores of “personal semantics” (e.g., Renoult, Davidson, Palombo, Moscovitch, & Levine, 2012).

In any one situation (or “event”), we activate schemas which can help us rapidly retrieve appropriate information as a scaffold (Irish & Piguët, 2013) but which also would create a metacognitive model of expectancies about fluency and the likelihood of retrieval, thus contributing to our evaluation of plausibility and a metacognitive representation of the situation or event.

The notion of a continuum should lead to some empirical tests of this idea, and the two domains may learn from each other. Most notably, the IAM literature has converged on a laboratory task which reliably yields IAMs. An obvious implication of our ideas is that in these mundane and repetitive tasks, participants should be experiencing metacognitive experiences and unresolved feelings of familiarity as well as more well-developed IAMs. Thus, when studying involuntary memories, one could ask for instances of déjà vu and familiarity to test some of the ideas suggested in the present paper. Also, questionnaire and diary studies should investigate the incidence of the two phenomena to examine whether those people who experience more IAMs also experience déjà vu more often.

Testing people with temporal-lobe pathology on tasks which generate IAMs would also be of interest (for a first case study for déjà vu, see Cleary et al., 2021b). Martin et al. (2012) have identified familiarity and recollection tasks which distinguish between people with temporal-lobe epilepsy who do and do not experience déjà vu, and according to our continuum hypothesis, people who experience more déjà vu in temporal-lobe epilepsy should also be more likely to experience involuntary memories. Such populations would help converge on the relationship between bottom-up and top-down processes in retrieval. Patients with selective parahippocampal impairments which lead to feeling less familiarity should, by our view, experience more spontaneous retrieval, and be less able to metacognitively act on cues in the environment. Such information is as yet unknown, primarily because, as with the existing IAM tasks in healthy populations, most memory measures concentrate on objective indices of declarative memory, and very little on subjective report or incomplete retrieval. Because we argue that familiarity operates to tell us that something in the environment “reminds” us of something, it should be possible to experimentally induce these feelings through priming and fluency manipulations, and measure the resultant effects on phenomenology and subjective experience.

To our knowledge, there are as yet no similar studies investigating involuntary memory retrieval in the context of brain damage, but there are studies demonstrating autobiographical memory retrieval deficits relating to clinical states (e.g., Múnica et al., 2014; Philippi, Tranel, Duff, & Rudrauf, 2013). We propose that it is a matter of time until case studies with people lacking involuntary retrieval will be reported. From the opposite end, there is already one case study by Parker, Cahill, and McGaugh (2006) describing AJ’s memory as “nonstop, uncontrollable, and automatic” which is an example of extreme domination of involuntary retrieval, described as most likely resulting from a variant of a neurodevelopmental frontostriatal disorder.

It would also be fruitful to study both the mental content and feeling of familiarity as two dimensions that may influence memory retrieval independently from each other. For example, one may try to manipulate familiarity and memory accessibility during involuntary retrieval by the use of different types of antecedent/previously prepared cue-sets such as, for example: (1) Cues that are highly efficient in quickly triggering an autobiographical memory and are also rated as highly familiar/typical,

(2) cues that are less efficient in triggering any concrete memory while being rated as familiar (i.e., that slowly but eventually leads to a memory retrieval), (3) cues that are less efficient and weakly familiar. As a result, one may expect that most IAMs will be experienced with the first type of cue, while déjà vu would be experienced with the second type.

Next, if we agree on the spontaneous nature of IAMs and déjà vu, then one may ask about the relationship between mind wandering and IAMs and déjà vu. As, in general, mind wandering relates to thoughts not related to the ongoing task per se, then one may expect to observe instances of IAMs and déjà vu during typical mind-wandering episodes. For instance, as Kvavilashvili et al. (2020; also Cole & Kvavilashvili, 2019; Plimpton et al., 2015) directly treat IAMs as an instance of spontaneous mind wandering, it may be argued that at least some examples of déjà vu may constitute the content of mind-wandering instances. Importantly, we propose that déjà vu and IAMs may both have multiple causes. For instance, it is still possible that on some occasions, déjà vu experiences are due to false recognition (as suggested by Mazzoni & Hanzakowski, 2011) or decoupled familiarity (Illman et al., 2012) on a neural level. It may also be true, especially when metacognitive abilities are impaired, that déjà vu is a result of the lack of access to already activated autobiographical memories that are below the awareness threshold (e.g., Conway & Loveday, 2015; Moulin et al., 2014).

Integrating déjà vu into autobiographical memory has opened interesting avenues for future research. For instance, recent advances and newer theoretical developments such as those proposed by Rubin and Umanath (2015; also Rubin, 2022, p. 10) propose a common theoretical ground and “home for many homeless categories of memory” including déjà vu. As they aim to improving and integrating theories arising from laboratory work, they create a conceptual “corner” space for such integration. Specifically, Rubin and Umanath (2015) proposed the theory of *event memory* to classify and understand “a mental construction of a scene recalled as a single occurrence” (p. 1).¹⁵ In these terms, déjà vu may be understood as an example of self-reference (i.e., about past events involving the person recalling them) implicit scene memory that “includes recognizing people and objects as a function of their context, traversing familiar routes without explicit memory” (Rubin, 2022, p. 7). Importantly, implicit scene memories are accompanied by feelings of familiarity and reliving without an access to memory of past experience. Although it is unknown whether déjà vu is associated with scene processing the same way as involuntary memories are, given the fact that scenes are known to be a common trigger of déjà vu (Brown, 2004; Cleary & Brown, 2022), it is highly possible that scenes, in general, may be a special form of triggering cues of both IAMs and déjà vu. This shows another possible way in which these two phenomena may have something in common and future studies should investigate this possibility.

Finally, we mentioned briefly several other memory-related phenomena (e.g., FOK, TOT, see Fig. 1) that strongly depend on the feeling of familiarity which were not discussed in the same detail as déjà vu and IAMs. Future research might want to consider how these primarily metacognitive experiences fit into a framework of thoughts and feelings associated with spontaneous retrieval. A key issue is their spontaneity – it would be normal that feelings of imminent retrieval and retrieval failure as well as feelings of familiarity guide retrieval processes, and our hypothesis here is that whilst we can ask participants to consciously and wilfully report their likelihood of retrieval in an

experimental task (e.g., in the FOK task), the TOT, the butcher on the bus, and FOK all also exist as naturally occurring spontaneous phenomena in daily life, and are far more frequently experienced than déjà vu.

4. Conclusions

In the present paper, we discussed the idea that IAMs and déjà vu can be described and explained as natural forms of memory processing and that both are based on the same retrieval processes. According to our continuum hypothesis, déjà vu is, at least on some occasions, an underdeveloped form and special case of involuntary memory; it draws upon the personal past and shares some crucial properties with IAMs, and surprisingly these two instances have not yet been comprehensively considered alongside each other in cognitive psychological discussions of human memory, in general, and autobiographical memory, in particular.

While there are a few possible origins of déjà vu and IAMs, not least in intercortical stimulation studies, we argue that one may be routine: A product of the ongoing automatic processing of cues in the environment which may present some conceptual or perceptual overlap with stored memory representations. As our cognitive system tries to match, as quickly and effortlessly as possible, the contents of mental representations stored in memory with the current contents of perception, one would expect déjà vu and involuntary memories to arise. This way, on some occasions, such an overlap may successfully trigger an involuntary memory of a given past personal event. This should be indeed expected from time to time, especially given the fact of “the vastness and richness of the autobiographical memory base” (Mace & Atkinson, 2009, p. 202). On some other occasions though, the overlap may be too weak and/or diffused to successfully trigger a given personal memory leaving one with only a feeling of familiarity, which we conceive as something like the TOT but for the personal past. Importantly, this should be also reasonably expected to happen given the fact that a great number of cues in the environment may coincide with a great number of memories stored in mind (i.e., the cue-overload principle) not being able to successfully match and trigger one given memory.

In terms of the function of these phenomena, we propose that they are the result of a continuously active memory system that automatically and rapidly scans the environment for matching representations. The purpose and evolutionary advantage of such a system is clear, quickly raising pertinent information to consciousness without effort. Yet, our characterisation is this system is of it being a constant iteration driven by feelings of familiarity and successful retrieval. Déjà vu arises when the familiarity arises in an inappropriate context, and its function is to signal that there is unwarranted activation of familiarity in the absence of retrieved contextual specifics which could link the present moment to the personal past.

Interestingly, there are neuropsychological patients which support this view. When familiarity is pathologically overactive, it can yield feelings of recollective confabulation, described by carers and medics as like permanent déjà vu (see Moulin, 2013, for a series of cases). The account of such patients is that rather than experiencing a conflict in mental evaluations as proposed here, they confabulate contextual specifics to justify the feeling of confabulation (Turner, Shores, Breen, & Coltheart, 2017). Neuroanatomically, in the most detailed analysis of the brain of such cases, Craik et al. (2014) describe extensive grey matter atrophy in frontal and medial temporal areas as shown by MRI, again

converging on the notion that inappropriate familiarity was not “discounted or edited out” (p. 367).

Here, we have rarely discussed conceptual knowledge in memory retrieval, focusing on the personal past. Déjà vu is studied exclusively from an episodic memory viewpoint, whereas involuntary memories are most studied in the context of a richer, more complex construct: Autobiographical memory encompasses both conceptual and event-specific information. As such, for example, whereas we might discuss the integration of episodic and semantic memory in autobiographical retrieval (e.g., Mace & Unlu, 2020), and can even propose purely semantic involuntary cognitions (e.g., Kvavilashvili & Mandler, 2004), there is not, to the best of our knowledge, a similar parallel in déjà vu: It seems to uniquely include an interpretation of a prior experience rather than information or knowledge. Although familiarity is an epistemic feeling which can be experienced for both semantic and episodic materials (see Kempnich, Urquhart, O'Connor, and Moulin [2017] for evidence of dual process retrieval processes in episodic word lists and general knowledge; or Bowles and Köhler [2014] for studies of familiarity in semantic memory), it is unclear what a “semantic” déjà vu would consist of. This issue still warrants attention in future work, although it would need to be discussed in the context of evidence shifting away from the existence of truly separable semantic and episodic systems (e.g., Renoult et al., 2019).

By relating to well-known and already elaborated mechanisms of autobiographical memory retrieval, we were able to explain the generation of déjà vu. This theoretical and conceptual integration will facilitate future studies on every-day spontaneous phenomena, and in doing so underlines the importance of three important dimensions on which memories are experienced: Intentionality, plausibility, and phenomenological experiences during retrieval, notably fluency.

Acknowledgements. We wish to thank five anonymous reviewers for their constructive, inspiring, and helpful comments on the earlier version of this manuscript.

Financial support. The paper was supported by the French Government Scholarship “Campus France” (K. B.); the National Science Centre, Poland (K. B., grant number UMO-2019/35/B/HS6/00528); the Bekker programme from the Polish National Agency for Academic Exchange (K. B., grant number PPN/BEK/2019/1/00092/DEC/1); and the Institut Universitaire de France (C. M.).

Competing interest. None.

Notes

1. IAMs have been differentiated from intrusive memories and flashbacks by Kvavilashvili (2014). She proposes a continuum with IAMs and flashbacks on opposite poles and intrusive memories in the middle. While they all share some features (all of them are spontaneously retrieved), they may also be treated as different. IAMs can be positive, negative or neutral, with non or minimal avoidance, disruption, distress; whereas intrusive memories can be positive or negative with moderate, high to very high avoidance, disruption and distress, and are typical for normal and clinical populations; while flashbacks that can only be negative with high to extreme avoidance, disruption and distress are restricted to the PTSD population only.
2. Inferior parietal lobule involvement could also be reflective of cognitive control driven by the frontoparietal control network.
3. The status of IAMs in people who are blind is unknown, but we would not expect this group to be unable to experience IAMs (see Tekcan et al., 2015, for differences in sighted and non-sighted participants' autobiographical memories).
4. This result was specific to older adults. The lifespan trajectory of IAMs and déjà vu is somewhat unknown, but interesting patterns with age emerge. There

are equivocal findings for IAMs, but the emerging view is that their frequency does not decline with age (e.g., Rubin & Berntsen, 2009), whereas the frequency of déjà vu does decrease with age (Moulin et al., 2014). IAMs are well documented in children as young as 35 months of age (e.g., Krøjgaard, Kingo, Jensen, & Berntsen, 2017; Sonne, Kingo, Berntsen, & Krøjgaard, 2019), but there are as yet no comparable studies for déjà vu. Since déjà vu can only be experienced by people with sufficient cognitive resources to oppose and detect the false feeling of familiarity, it may not be experienced by children and individuals without sufficient metacognitive function. It is relatively easy to imagine that a child could more readily report a memory coming to mind spontaneously than the unusual feelings of familiarity in déjà vu, but this is an area which needs more research.

5. Familiarity is not only thought to be strictly perceptual in nature but may also be described as conceptual (e.g., Rajaram & Geraci, 2000).

6. Fluency is not thought to be domain specific, that is, reserved to evaluations about memory, but is a heuristic which can be applied to many contexts and situations (a “*heuristic inference processes*”; e.g., Mantonakis and Hastie, 2011). Depending on what is asked of the participant, fluency can act on judgements of memory, preference, confidence, intentionality and so on.

7. Many of these works draw inspiration from Norman & Bobrow’s (1979, p. 109) paper where they posed the questions: “How does one know what is needed from memory? Is not the knowledge of what is sought in itself part of the knowledge that is being sought?”

8. This is not to say that a reciprocal feedback loop is not possible. It may be that initially familiarity prompts an attempt to retrieve contextual specifics. Once such an attempt fails, it may be felt even more strongly, leading to even a more intense and goal-oriented search.

9. The idea that uniqueness of cues plays an important role in the memory retrieval is not so new (e.g., Nairne, 2002; Poirier et al., 2012), and it may be considered a common notion in memory research. However, as we argue in the present paper, the novelty of our account lies in the idea that although a typical and common cue may not be sufficient to trigger a particular memory (as it is in case of unique cues), they still may activate and trigger an unidentified feeling of familiarity leading to a déjà vu experience.

10. Cue familiarity is an experimental variable that has drawn a lot of attention in experiments on recognition memory (e.g., Ryals & Cleary, 2012) and metacognition (e.g., Metcalfe, Schwartz, & Joaquim, 1993). It is used in a diverse set of experimental contexts including “don’t know” responding (e.g., Hanczakowski, Pasek, Zawadzka, & Mazzoni, 2013); illusory recollective experience (e.g., Huebert et al., 2022); decision making (e.g., Schwikert & Curran, 2014) and autobiographical memory (e.g., Fenerci & Sheldon, 2022). Much of the ideas presented here draw upon these literatures, and in particular the series of experiments by Cleary and colleagues on déjà vu reviewed in this article.

11. Although these studies suggest that the retrieval of IAMs may not depend exclusively on inhibitory control, there is still a possibility that such a suppression mechanism may not to be switched on all the time, but is only switched on intermittently, similarly to proactive and retroactive cognitive control (for a review, see Braver, 2012). However, there is still the need to find circumstances in which such a mechanism may be observed. As no adequate answers to these questions have been found yet, our position is relatively inclusive; namely, at this stage, we do not say categorically that the cognitive inhibitory account should be disregarded completely.

12. These stages are not necessarily consecutive, where one stage is terminated before another is activated. While such a simple presentation facilitates our understanding of memory retrieval, they are rather thought of as dynamic, complex floating system of memory states where all of the stages may be activated simultaneously to different extents at the same time. Thus, we can think about these stages in a similar way as we understand memory circuitry between the temporal and frontal lobes.

13. Note that while the idea that memory retrieval is or is not an all-or-none process is broadly discussed in the memory literature (e.g., Kempnich et al., 2017; Onyper, Zhang, & Howard, 2010; Wixted & Mickes, 2010; Vilberg & Rugg, 2008; Yonelinas, 2002), in general, this idea is rarely (if at all) discussed within the field of autobiographical memory retrieval.

14. One interesting hypothesis to fall out of this proposal is that older adults’ lack of déjà vu experiences could be explained by changes in how plausible subsequent retrieval is. That is, when confronted with retrieval failure and a

feeling of familiarity, the older adult would interpret this as a result of temporary retrieval failure and that later retrieval was possible, and the context for memory retrieval was plausible. As a result, one prediction is that as feelings of déjà vu diminish, feelings of FOK/TOT should increase.

15. Maguire and Mullally (2013) contributed to the notion that scene construction may serve a special purpose in episodic and autobiographical memory and their work likely prompted subsequent research on that idea; namely, the role of scene processing in episodic and autobiographical memory retrieval.

References

- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063–1087. <https://doi.org/10.1037//0278-7393.20.5.1063>
- Andrews-Hanna, J. R., Reidler, J. S., Huang, C., & Buckner, R. L. (2010). Evidence for the default network’s role in spontaneous cognition. *Journal of Neurophysiology*, 104, 322–335.
- Arango-Muñoz, S. (2011). Two levels of metacognition. *Philosophia*, 39, 71–82.
- Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology*, 27(2), 193–211. doi:10.1080/09515089.2012.73200
- Ball, C. T. (2007). Can we elicit involuntary autobiographical memories in the laboratory? In J. H. Mace (Ed.), *Involuntary memory* (pp. 127–152). Blackwell.
- Ball, C. T., & Hennessey, J. (2009). Subliminal priming of autobiographical memories. *Memory*, 17(3), 311–322. doi:10.1080/09658210902729483
- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74(2), 81.
- Barzykowski, K. (2014). How can we catch spontaneous memories: A review of methodological issues in involuntary autobiographical memories studies. SAGE research methods cases. London: SAGE Publications, Ltd. <https://doi.org/10.4135/978144627305013517801>
- Barzykowski, K., Hajdas, S., Radel, R., & Kvavilashvili, L. (2022). Effects of inhibitory control capacity and cognitive load on involuntary past and future thoughts: A laboratory study. *Consciousness and Cognition*, 102, 1–19. <https://doi.org/10.1016/j.concog.2022.103353>
- Barzykowski, K., Hajdas, S., Radel, R., Niedźwieńska, A., & Kvavilashvili, L. (2021a). The role of inhibitory control and ADHD symptoms in the occurrence of involuntary thoughts about the past and future: An individual differences study. *Consciousness and Cognition*, 95, 1–19. doi:10.1016/j.concog.2021.103208
- Barzykowski, K., & Mazzoni, G. (2022). Do intuitive ideas of the qualities that should characterize involuntary and voluntary memories affect their classification? *Psychological Research*, 86, 170–195. doi:10.1007/s00426-020-01465-3
- Barzykowski, K., & Niedźwieńska, A. (2012). Przegląd badań nad mimowolnymi wspomnieniami autobiograficznymi. Perspektywy badawcze. [Review of studies on involuntary autobiographical memories. Research perspectives]. *Roczniki Psychologiczne [Annals of Psychology]*, 1(XV), 55–74.
- Barzykowski, K., & Niedźwieńska, A. (2016). The effects of instruction on the frequency and characteristics of involuntary autobiographical memories. *PLoS ONE*, 11(6), 1–15. <http://dx.doi.org/10.1371/journal.pone.0157121>
- Barzykowski, K., & Niedźwieńska, A. (2018a). Involuntary autobiographical memories are relatively more often reported during high cognitive load tasks. *Acta Psychologica*, 182, 119–128. doi:10.1016/j.actpsy.2017.11.014
- Barzykowski, K., & Niedźwieńska, A. (2018b). Priming involuntary autobiographical memories in the lab. *Memory*, 26(2), 277–289. doi:10.1080/09658211.2017.1353102
- Barzykowski, K., Niedźwieńska, A., & Mazzoni, G. (2019a). How intention to retrieve a memory and expectation that it will happen influence retrieval of autobiographical memories. *Consciousness and Cognition*, 72, 31–48. doi:10.1016/j.concog.2019.03.011
- Barzykowski, K., Radel, R., Niedźwieńska, A., & Kvavilashvili, L. (2019b). Why are we not flooded by involuntary thoughts about past and future? Testing the cognitive inhibition dependency hypothesis. *Psychological Research*, 83(4), 666–683. <https://doi.org/10.1007/s00426-018-1120-6>
- Barzykowski, K., Skopicz-Radkiewicz, E., Kabut, R., Staugaard, S. R., & Mazzoni, G. (2023). Intention and monitoring influence the content of memory reports. *Psychological Reports*, 126(2), 918–945. <https://doi.org/10.1177/00332941211048736>
- Barzykowski, K., & Staugaard, S. R. (2016). Does retrieval intentionality really matter? Similarities and differences between involuntary memories and directly and generatively retrieved voluntary memories. *British Journal of Psychology*, 107(3), 519–536. doi:10.1111/bjop.12160
- Barzykowski, K., & Staugaard, S. R. (2018). How intention and monitoring your thoughts influence characteristics of autobiographical memories. *British Journal of Psychology*, 109(2), 321–340. doi:10.1111/bjop.12259
- Barzykowski, K., Staugaard, S. R., & Mazzoni, G. (2021b). Retrieval effort or intention: Which is more important for participants’ classification of involuntary and voluntary memories? *British Journal of Psychology*, 112(4), 1080–1102. doi:10.1111/bjop.12498

- Bastin, C., Besson, G., Simon, J., Delhay, E., Geurten, M., Willems, S., & Salmon, E. (2019). An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, 42, e281: 1–60. doi:10.1017/S0140525X19000621
- Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology*, 10, 435–454. doi:10.1002/(SICI)1099-0720(199610)10:5<435::AID-ACP408>3.0.CO;2-L
- Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, 6, 113–141.
- Berntsen, D. (2009). *Involuntary autobiographical memories. An introduction to the unbidden past*. Cambridge University Press.
- Berntsen, D. (2010). The unbidden past: Involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science*, 19(3), 138–142.
- Berntsen, D. (2015). From everyday life to trauma: Research on everyday involuntary memories advances our understanding of intrusive memories of trauma. In L. Watson, & D. Berntsen (Eds.), *Clinical perspectives on autobiographical memory* (pp. 172–196). Cambridge University Press.
- Berntsen, D., Rubin, D. C., & Salgado, S. (2015). The frequency of involuntary autobiographical memories and future thoughts in relation to daydreaming, emotional distress, and age. *Consciousness and Cognition*, 36, 352–372.
- Berntsen, D., Staagaard, S. R., & Sørensen, L. M. T. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General*, 142(2), 426–444. doi:10.1037/a0029128
- Berry, C. J., Shanks, D. R., Speekenbrink, M., & Henson, R. N. (2012). Models of recognition, repetition priming, and fluency: Exploring a new framework. *Psychological Review*, 119(1), 40–79.
- Besson, G., Ceccaldi, M., Didic, M., & Barbeau, E. J. (2012). The speed of visual recognition memory. *Visual Cognition*, 20(10), 1131–1152.
- Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford University Press.
- Botzung, A., Denkova, E., Ciuciu, P., Scheiber, C., & Manning, L. (2008). The neural bases of the constructive nature of autobiographical memories studied with a self-paced fMRI design. *Memory*, 16, 351–363. doi:10.1080/09658210801931222
- Bowles, B., & Köhler, S. (2014). Availability of semantic knowledge in familiar-only experiences for names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3), 724–737. https://doi.org/10.1037/a0035584
- Bradley, R. J., Moulin, C. J. A., & Kvavilashvili, L. (2013). Involuntary autobiographical memories. *The Psychologist*, 26(3), 190–193.
- Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, 16(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010
- Brewin, C. R., Gregory, J. D., Lipton, M., & Burgess, N. (2010). Intrusive images in psychological disorders: Characteristics, neural mechanisms, and treatment implications. *Psychological Review*, 117, 210–232.
- Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109(2), 204–223. https://doi.org/10.1037/0033-2909.109.2.204
- Brown, A. S. (2003). A review of the déjà vu experience. *Psychological Bulletin*, 129, 394–413.
- Brown, A. S. (2004). *The déjà vu experience*. Psychology Press.
- Brown, A. S., & Marsh, E. J. (2008). Evoking false beliefs about autobiographical experience. *Psychonomic Bulletin and Review*, 15, 186–190.
- Brown, A. S., & Marsh, E. J. (2009). Creating illusions of past encounter through brief exposure. *Psychological Science*, 20, 534–538.
- Brown, A. S., & Marsh, E. J. (2010). Digging into déjà vu: Recent research on possible mechanisms. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 53, pp. 33–62). Academic Press.
- Brunec, I. K., Bellana, B., Ozubko, J. D., Man, V., Robin, J., Liu, Z. X., ... Moscovitch, M. (2018). Multiple scales of representation along the hippocampal anteroposterior axis in humans. *Current Biology*, 28(13), 2129–2135.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124, 1–38. https://doi.org/10.1196/annals.1440.011
- Burgess, P. W., & Shallice, T. (1996). Confabulation and the control of recollection. *Memory*, 4, 359–411.
- Cai, W., Chen, T., Ryals, S., Kochalka, J., Li, C. S. R., & Menon, V. (2016). Causal interactions within a frontal-cingulate-parietal network during cognitive control: Convergent evidence from a multisite-multitask investigation. *Cerebral Cortex*, 26(5), 2140–2153.
- Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, N. R., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Review Neuroscience*, 17, 718–731. https://doi.org/10.1038/nrn.2016.113
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62, 73–101.
- Clark, I. A., Mackay, C. E., & Holmes, E. A. (2013). Positive involuntary autobiographical memories: You first have to live them. *Consciousness and Cognition*, 22, 402–406. https://doi.org/10.1016/j.concog.2013.01.008
- Cleary, A. M. (2008). Recognition memory, familiarity, and déjà vu experiences. *Current Directions in Psychological Science*, 17(5), 353–357.
- Cleary, A. M., & Brown, A. S. (2022). *The Déjà vu Experience* (2nd ed.). Routledge.
- Cleary, A. M., Brown, A. S., Sawyer, B. D., Nomi, J. S., Ajoku, A. C., & Ryals, A. J. (2012). Familiarity from the configuration of objects in 3-dimensional space and its relation to déjà vu: A virtual reality investigation. *Consciousness and Cognition*, 21(2), 969–975.
- Cleary, A. M., & Claxton, A. B. (2018). Déjà vu: An illusion of prediction. *Psychological Science*, 29(4), 635–644.
- Cleary, A. M., Huebert, A. M., McNeely-White, K. L., & Spahr, K. S. (2019). A postdictive bias associated with déjà vu. *Psychonomic Bulletin & Review*, 26(4), 1433–1439.
- Cleary, A. M., McNeely-White, K. L., Huebert, A. M., & Claxton, A. B. (2021a). Déjà vu and the feeling of prediction: An association with familiarity strength. *Memory*, 29(7), 904–920.
- Cleary, A. M., Morris, A. L., & Langley, M. M. (2007). Recognition memory for novel stimuli: The structural regularity hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2), 379–393.
- Cleary, A. M., Neisser, J., McMahan, T., Parsons, T. D., Alwaki, A., Okada, N., ... Pedersen, N. P. (2021b). Subjective distinguishability of seizure and non-seizure Déjà Vu: A case report, brief literature review, and research prospects. *Epilepsy & Behavior*, 125, 108373.
- Cleary, A. M., Ryals, A. J., & Nomi, J. S. (2009). Can déjà vu result from similarity to a prior experience? Support for the similarity hypothesis of déjà vu. *Psychonomic Bulletin & Review*, 16, 1082–1088.
- Cole, S., & Kvavilashvili, L. (2019). Spontaneous future cognition: The past, present and future of an emerging topic. *Psychological Research*, 83(4), 631–650. https://doi.org/10.1007/s00426-019-01193-3
- Cole, S., & Kvavilashvili, L. (2021). Spontaneous and deliberate future thinking: A dual process account. *Psychological Research*, 85(2), 464–479. https://doi.org/10.1007/s00426-019-01262-7
- Conway, M. A. (1990). Associations between autobiographical memories and concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 799–812.
- Conway, M. A. (2001). Sensory-perceptual episodic memory and its context: Autobiographical memory. Philosophical transactions of the Royal Society of London. Series B: Biological Sciences, 356, 1375–1384. doi:10.1098/rstb.2001.0940
- Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, 53, 594–628.
- Conway, M. A. (2008). Exploring episodic memory. In E. Dere, A. Easton, L. Nadel, & J. P. Huston (Eds.), *Handbook of episodic memory* (Vol. 18, pp. 19–29). Elsevier.
- Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, 47(11), 2305–2313. doi:10.1016/j.neuropsychologia.2009.02.003
- Conway, M. A., & Jobson, L. (2012). On the nature of autobiographical memory. In D. Berntsen, & D. C. Rubin (Eds.), *Understanding autobiographical memory. Theories and approaches* (pp. 54–69). Cambridge University Press.
- Conway, M. A., & Loveday, C. (2010). Accessing autobiographical memories. In J. H. Mace (Ed.), *The act of remembering. Toward an understanding of how we recall the past* (pp. 56–70). Wiley-Blackwell.
- Conway, M. A., & Loveday, C. (2015). Remembering, imagining, false memories & personal meanings. *Consciousness and Cognition*, 33, 574–581. http://dx.doi.org/10.1016/j.concog.2014.12.002
- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107, 261–288. doi:10.1037/0033-295X.107.2.261
- Craik, F. I., Barene, M. D., Rathbone, C. J., Grusec, J. E., Stuss, D. T., Gao, F., ... Black, S. E. (2014). VL: A further case of erroneous recollection. *Neuropsychologia*, 56, 367–380.
- Curot, J., Busigny, T., Valton, L., Denuelle, M., Vignat, J.-P., Maillard, L., ... Barbeau, E. J. (2017). Memory scrutinized through electrical brain stimulation: A review of 80 years of experiential phenomena. *Neuroscience and Biobehavioral Reviews*, 78, 161–177. https://doi.org/10.1016/j.neubiorev.2017.04.018
- Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia*, 37(7), 771–785. http://doi.org/10.1016/S0028-3932(98)00133-X
- Davachi, L., & Dobbins, I. G. (2008). Declarative memory. *Current Directions in Psychological Science*, 17, 112–118. https://doi.org/10.1111/j.1467-8721.2008.00559.x
- De Brigard, F., Spreng, R. N., Mitchell, J. P., & Schacter, D. L. (2015). Neural activity associated with self, other, and object-based counterfactual thinking. *NeuroImage*, 109, 12–26.
- Delay, C. G., & Wixted, J. T. (2021). Discrete-state versus continuous models of the confidence-accuracy relationship in recognition memory. *Psychonomic Bulletin & Review*, 28(2), 556–564.
- de Sousa, R. (2009). Epistemic feelings. *Mind & Matter*, 7, 139–161.
- Dewhurst, S. A., Holmes, S. J., Brandt, K. R., & Dean, G. M. (2006). Measuring the speed of the conscious components of recognition memory: Remembering is faster than knowing. *Cognition and Consciousness*, 15, 673–686.
- Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, 111(2), 524–542.

- Duzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H.-J., & Tulving, E. (1997). Event-related potential correlates of two states of conscious awareness in memory. *Proceedings of The National Academy of Sciences*, *94*, 5973–5978.
- Eichenbaum, H. (2017). Memory: Organization and control. *Annual Review of Psychology*, *68*, 19–45.
- Fenerci, C., & Sheldon, S. (2022). The role of episodic memory in imagining autobiographical events: The influence of event expectancy and context familiarity. *Memory*, *30*(5), 573–590. <https://doi.org/10.1080/09658211.2022.2032178>
- Gerrans, P. (2014). Pathologies of hyper familiarity in dreams, delusions and déjà vu. *Frontiers in Psychology*, *5*, 97.
- Gilbert, S. J., Hadjipavlou, N., & Raelison, M. (2013). Automaticity and control in prospective memory: A computational model. *PLoS ONE*, *8*(3), e59852. <https://doi.org/10.1371/journal.pone.0059852>
- Goldinger, S. D., & Hansen, W. A. (2005). Remembering by the seat of your pants. *Psychological Science*, *16*, 525–529.
- Goldsmith, M., & Koriat, A. (1999). The strategic regulation of memory reporting: Mechanisms and performance consequences. In D. Gopher, & A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 373–400). MIT Press.
- Greely, N. (2021). Epistemic feelings, metacognition, and the Lima problem. *Synthese*, *199*, 6803–6825. <https://doi.org/10.1007/s11229-021-03094-8>
- Hall, N. M., Gjedde, A., & Kupers, R. (2008). Neural mechanisms of voluntary and involuntary recall: A PET study. *Behavioural Brain Research*, *186*, 261–272. doi:10.1016/j.bbr.2007.08.026
- Hall, S. A., Rubin, D. C., Miles, A., Davis, S. W., Wing, E. A., Cabeza, R., & Berntsen, D. (2014). The neural basis of involuntary episodic memories. *Journal of Cognitive Neuroscience*, *26*(10), 2385–2399. https://doi.org/10.1162/jocn_a_00633
- Hanczakowski, M., Pasek, T., Zawadzka, K., & Mazzoni, G. (2013). Cue familiarity and “don’t know” responding in episodic memory tasks. *Journal of Memory and Language*, *69*(3), 368–383.
- Harris, C. B., O’Connor, A. R., & Sutton, J. (2015). Cue generation and memory construction in direct and generative autobiographical memory retrieval. *Consciousness and Cognition*, *33*, 204–216. doi:10.1016/j.concog.2014.12.012
- Hawthorne, N. (1863). *Our old home: A series of English sketches*. The Floating Press.
- Heine, M. K., Ober, B. A., & Shenaut, G. K. (1999). Naturally occurring and experimentally induced tip-of-the-tongue experiences in three adult age groups. *Psychology and Aging*, *14*, 445–457.
- Hintzman, D. L., Caulton, D. A., & Levitin, D. J. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory & Cognition*, *26*(3), 449–462.
- Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence of separate processes of familiarity and recall. *Journal of Memory and Language*, *33*, 1–18.
- Huebert, A. M., McNeely-White, K. L., & Cleary, A. M. (2022). Can cue familiarity during recall failure prompt illusory recollective experience? *Memory & Cognition*, *50*(4), 681–695. <https://doi.org/10.3758/s13421-021-01248-0>
- Hyman, I. E., Jr., Burland, N. K., Duskin, H. M., Cook, M. C., Roy, C. M., McGrath, J. C., & Roundhill, R. F. (2013). Going gaga: Investigating, creating, and manipulating the song stuck in my head. *Applied Cognitive Psychology*, *27*, 204–215. <https://doi.org/10.1002/acp.2897>
- Illman, N. A., Butler, C. R., Souchay, C., & Moulin, C. J. A. (2012). Déjà vu experiences in temporal lobe epilepsy. *Epilepsy Research and Treatment*, 1–15. doi: 10.1155/2012/539567
- Irish, M., & Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. *Frontiers in Behavioral Neuroscience*, *7*, 27.
- Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, *32*, 43–49.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340. doi:10.1037/0096-3445.110.3.306
- Jersakova, R., O’Connor, A. R., & Moulin, C. (2015). What’s new in déjà vu? In S. Haque & E. Sheppard (Eds.), *Culture and cognition: A collection of critical essays* (pp. 137–150). Peter Lang.
- Johannessen, K. B., & Berntsen, D. (2010). Current concerns in involuntary and voluntary autobiographical memories. *Consciousness and Cognition*, *19*, 847–860. doi:10.1016/j.concog.2010.01.009
- Karpicke, J. D., McCabe, D. P., & Roediger, H. L. (2008). False memories are not surprising: The subjective experience of an associative memory illusion. *Journal of Memory and Language*, *58*, 1065–1079.
- Kempnich, M., Urquhart, J. A., O’Connor, A. R., & Moulin, C. J. A. (2017). Evidence for the contribution of a threshold retrieval process to semantic memory. *Quarterly Journal of Experimental Psychology*, *70*(10), 2026–2047. <https://doi.org/10.1080/17470218.2016.1220607>
- Klein, S. B. (2013). The temporal orientation of memory: It’s time for a change of direction. *Journal of Applied Research in Memory and Cognition*, *2*(4), 222–234. doi:10.1016/j.jarmac.2013.08.001
- Knight, E. F. (1895). *Where three empires meet: A narrative of recent travel in Kashmir, Western Tibet, Gilgit and the adjoining countries*. Longmans, Green & Co.
- Kolers, P. A., & Palef, S. R. (1976). Knowing not. *Memory & Cognition*, *4*, 553–558.
- Kompus, K., Eichele, T., Hugdahl, K., & Nyberg, L. (2011). Multimodal imaging of incidental retrieval: The low route to memory. *Journal of Cognitive Neuroscience*, *23*(4), 947–960. <https://doi.org/10.1162/jocn.2010.21494>
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*(3), 490–517. <https://doi.org/10.1037/0033-295X.103.3.490>
- Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings-of-knowing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 34–53.
- Krojjgaard, P., Kingo, O. S., Jensen, T. S., & Berntsen, D. (2017). By-passing strategic retrieval: Experimentally induced spontaneous episodic memories in 35- and 46-month-old children. *Consciousness and Cognition*, *55*, 91–105. <https://doi.org/10.1016/j.concog.2017.08.001>
- Kuhl, B. A., & Wagner, A. D. (2009). Forgetting and retrieval. In G. G. Berntson, & J. T. Cacioppo (Eds.), *Handbook of neuroscience for the behavioral sciences* (Vol. 1, pp. 586–605). John Wiley & Sons Inc. <https://doi.org/10.1002/9780470478509.neubb001031>
- Kvavilashvili, L. (2014). Solving the mystery of intrusive flashbacks in posttraumatic stress disorder: Comment on Brewin (2014). *Psychological Bulletin*, *140*(1), 98–104. doi:10.1037/a0034677
- Kvavilashvili, L., & Mandler, G. (2004). Out of one’s mind: A study of involuntary semantic memories. *Cognitive Psychology*, *48*, 47–94.
- Kvavilashvili, L., Niedźwieńska, A., Gilbert, S. J., & Markostamou, I. (2020). Deficits in spontaneous cognition as an early marker of Alzheimer’s disease. *Trends in Cognitive Sciences*, *24*(4), 285–301. <https://doi.org/10.1016/j.tics.2020.01.005>
- Kvavilashvili, L., & Rummel, J. (2020). On the nature of everyday prospection: A review and theoretical integration of research on mind-wandering, future thinking, and prospective memory. *Review of General Psychology*, *24*, 210–237. <https://doi.org/10.1177/1089268020918843>
- Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, *14*(4), 334–339.
- Laughland, A., & Kvavilashvili, L. (2018). Should participants be left to their own devices? Comparing paper and smartphone diary diaries in psychological research. *Journal of Applied Research in Memory and Cognition*, *7*, 552–563.
- Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract cuing: The Proustian view is incorrect. *Applied Cognitive Psychology*, *18*, 893–899. doi:10.1002/acp.1020
- Mace, J. H. (2005). Priming involuntary autobiographical memories. *Memory*, *13*, 874–884.
- Mace, J. H. (2006). Episodic remembering creates access to involuntary conscious memory: Demonstrating involuntary recall on a voluntary recall task. *Memory*, *14*, 917–924. <https://doi.org/10.1080/09658210600759766>
- Mace, J. H. (2007). Does involuntary remembering occur during voluntary remembering? In J. H. Mace (Ed.), *Involuntary memory* (pp. 50–67). Blackwell.
- Mace, J. H., & Atkinson, E. (2009). Can we determine the functions of everyday involuntary autobiographical memories? In M. R. Kelley (Ed.), *Applied memory* (pp. 199–212). Nova Science Publishers, Inc.
- Mace, J. H., Clevinger, A. M., Delaney, D. M., Mendez, A. S., & Simpson, S. H. (2017). Voluntary remembering: Elucidating the mental strategies used to recall the past. *Applied Cognitive Psychology*, *31*, 156–163. doi:10.1002/acp.3313
- Mace, J. H., & Unlu, M. (2020). Semantic-to-autobiographical memory priming occurs across multiple sources: Implications for autobiographical remembering. *Memory & Cognition*, *48*(6), 931–941. <https://doi.org/10.3758/s13421-020-01029-1>
- Maguire, E. A., & Mullally, S. L. (2013). The hippocampus: A manifesto for change. *Journal of Experimental Psychology: General*, *142*(4), 1180–1189. <https://doi.org/10.1037/a0033650>
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271.
- Mandler, G. (2008). Familiarity breeds attempts. *Perspectives on Psychological Science*, *3*, 390–399.
- Mantonakis, A., & Hastie, R. (2011). Surprising fluency: Bruce Whittlesea’s contributions to our understanding of the role of fundamental adaptive cognitive processes. In P. A. Higham, & J. P. Leboe (Eds.), *Constructions of remembering and metacognition: Essays in honor of Bruce Whittlesea* (pp. 201–214). Palgrave Macmillan.
- Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langa, G., ... Smallwood, J. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, *113* (44), 12574–12579.
- Marks, E. H., Franklin, A. R., & Zoellner, L. A. (2018). Can’t get it out of my mind: A systematic review of predictors of intrusive memories of distressing events. *Psychological Bulletin*, *144*(6), 584–640. <https://doi.org/10.1037/bul0000132>
- Marshall, J. C., Halligan, P. W., & Wade, D. T. (1995). Reduplication of an event after head-injury: A cautionary case-report. *Cortex*, *31*, 183–190.

- Martin, C. B., Mirsattari, S. M., Pruessner, J. C., Pietrantonio, S., Burneo, J. G., Hayman-Abello, B., & Köhler, S. (2012). Déjà vu in unilateral temporal-lobe epilepsy is associated with selective familiarity impairments on experimental tasks of recognition memory. *Neuropsychologia*, *50*(13), 2981–2991.
- Mazzoni, G., & Hanczakowski, M. (2011). Metacognitive processes before and during retrieval. In P. A. Higham, & J. P. Leboe (Eds.), *Constructions of remembering and metacognition: Essays in honor of Bruce Whittlesea* (pp. 91–106). Palgrave Macmillan.
- Mazzoni, G., & Kirsch, I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model. In T. J. Perfect, & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 121–145). Cambridge University Press.
- McDermott, K. B., Gilmore, A. W., Nelson, S. M., Watson, J. M., & Ojemann, J. G. (2017). The parietal memory network activates similarly for true and associative false recognition elicited via the DRM procedure. *Cortex*, *87*, 96–107. doi:10.1016/j.cortex.2016.09.008
- Menon, V., & D'Esposito, M. (2022). The role of PFC networks in cognitive control and executive function. *Neuropsychopharmacology*, *47*(1), 90–103.
- Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 851.
- Miller, G. A. (1962/1974). *Psychology: The science of mental life*. Penguin.
- Mills, C., Zamani, A., White, R., & Christoff, K. (2021). Out of the blue: Understanding abrupt and wayward transitions in thought using probability and predictive processing. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, *376*(1817), 20190692.
- Milton, F., Butler, C. R., & Zeman, A. Z. (2011). Transient epileptic amnesia: Déjà vu heralding recovery of lost memories. *Journal of Neurology, Neurosurgery & Psychiatry*, *82*, 1178–1179.
- Montaldi, D., & Mayes, A. R. (2010). The role of recollection and familiarity in the functional differentiation of the medial temporal lobes. *Hippocampus*, *20*(11), 1291–1314. <https://doi.org/10.1002/hipo.20853>
- Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *Journal of Neuroscience*, *38*(14), 3534–3546.
- Moscovitch, M. (2008). The hippocampus as a “stupid,” domain-specific module: Implications for theories of recent and remote memory, and of imagination. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *62*(1), 62–79. <https://doi.org/10.1037/1196-1961.62.1.62>
- Moulds, M. L., & Krans, J. (2015). Intrusive, involuntary memories in depression. In L. Watson, & D. Bernstein (Eds.), *Clinical perspectives on autobiographical memory* (pp. 154–171). Cambridge University Press.
- Moulin, C. J. (2013). Disordered recognition memory: Recollective confabulation. *Cortex*, *49*(6), 1541–1552.
- Moulin, C. J. A., & Souchay, C. (2014). Epistemic feelings and memory. In T. Perfect & D. S. Lindsay (Eds.), *The SAGE handbook of applied memory* (pp. 520–538). SAGE Publications. <https://doi.org/10.4135/9781446294703.n29>
- Moulin, C. J. A. (2018). *The cognitive neuropsychology of déjà vu. Essays in cognitive psychology*. Psychology Press (Routledge).
- Moulin, C. J. A., & Souchay, C. (2014). Epistemic feelings and memory. In T. Perfect, & S. Lindsay (Eds.), *Sage handbook of applied memory* (pp. 520–539). Sage Publications.
- Moulin, C. J. A., Souchay, C., Bradley, R., Buchanan, S., Karadoller, D. Z., & Akan, M. (2014). Déjà vu in older adults. In B. Schwartz, & A. Brown (Eds.), *The tip of the tongue states and related phenomena* (pp. 281–304). Cambridge University Press.
- Moulin, C. J. A., Souchay, C., & Morris, R. G. (2013). The cognitive neuropsychology of recollection. *Cortex*, *9*, 1445–1451.
- Múnera, C. P., Lomlomdjan, C., Gori, B., Terpiluk, V., Medel, N., Solís, P., & Kochen, S. (2014). Episodic and semantic autobiographical memory in temporal lobe epilepsy. *Epilepsy Research and Treatment*, 1–9. <https://doi.org/10.1155/2014/157452>
- Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory*, *10*, 389–395.
- Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review*, *19*, 969–974. doi:10.3758/s13423-012-0292-0
- Niedźwieńska, A., & Barzykowski, K. (2012). The age prospective memory paradox within the same sample in time-based and event-based tasks. *Aging, Neuropsychology, and Cognition*, *19*(1–2), 58–83. doi:10.1080/13825585.2011.628374, IF: 1.919.
- Niedźwieńska, A., Rendell, P., Barzykowski, K., & Leszczyńska, A. (2014). Only social feedback reduces age-related prospective memory deficits in “virtual week”. *International Psychogeriatrics*, *26*(5), 759–767. doi:10.1017/S1041610214000027
- Norman, D. A., & Bobrow, D. G. (1979). An intermediate stage in memory retrieval. *Cognitive Psychology*, *1*, 107–123.
- O'Connor, A. R., Lever, C., & Moulin, C. J. A. (2010). Novel insights into false recollection: A model of déjà vu. *Cognitive Neuropsychiatry*, *15*, 118–144.
- O'Connor, A. R., & Moulin, C. J. A. (2010). Recognition without identification, erroneous familiarity, and déjà vu. *Current Psychiatry Reports*, *12*(3), 165–173. <http://doi.org/10.1007/s11920-010-0119-5>
- O'Connor, A. R., & Moulin, C. J. A. (2006). Normal patterns of déjà vu experience in a healthy, blind male: Challenging optical pathway delay theory. *Brain and Cognition*, *62*, 246–249.
- Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *Journal of Experimental Psychology: General*, *139*(2), 341–364. <https://doi.org/10.1037/a0018926>
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776.
- Ortiz De Gortari, A. B., & Griffiths, M. D. (2015). Game transfer phenomena and its associated factors: An exploratory empirical online survey study. *Computers in Human Behavior*, *51*(PA), 195–202. <https://doi.org/10.1016/j.chb.2015.04.060>
- Parker, E. S., Cahill, L., & McGaugh, J. L. (2006). A case of unusual autobiographical remembering. *Neurocase: Case Studies in Neuropsychology, Neuropsychiatry, and Behavioural Neurology*, *12*(1), 35–49. <https://doi.org/10.1080/13554790500473680>
- Penfield, W., & Perot, P. (1963). The brain's record of auditory and visual experience: A final summary and discussion. *Brain*, *86*, 595–696.
- Perrin, D., & Rousset, S. (2014). The episodicity of memory. *Review of Philosophy and Psychology*, *5*(3), 291–312.
- Philippi, C. L., Tranel, D., Duff, M., & Rudrauf, D. (2013). Damage to the default mode network disrupts autobiographical memory retrieval. *Social Cognitive and Affective Neuroscience*, *10*(3), 318–326. <https://doi.org/10.1093/scan/nsu070>
- Plimpton, B., Patel, P., & Kvavilashvili, L. (2015). Role of triggers and dysphoria in mind-wandering about past, present and future: A laboratory study. *Consciousness and Cognition*, *33*, 261–276.
- Poirier, M., Nairne, J. S., Morin, C., Zimmerman, F. G. S., Koutmeridou, K., & Fowler, J. (2012). Memory as discrimination: A challenge to the encoding-retrieval match principle. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*, 16–29.
- Quamme, J. R., Yonelinas, A. P., & Norman, K. A. (2007). Effect of unitization on associative recognition in amnesia. *Hippocampus*, *17*(3), 192–200. <https://doi.org/10.1002/hipo.20257>
- Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*, 433–447.
- Rajaram, S., & Geraci, L. (2000). Conceptual fluency selectively influences knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(4), 1070–1074.
- Rasmussen, A. S., & Berntsen, D. (2011). The unpredictable past: Spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition*, *20*(4), 1842–1846. <https://doi.org/10.1016/j.concog.2011.07.010>
- Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, *19*, 90–138.
- Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 435–451.
- Renoult, L., Davidson, P. S., Palombo, D. J., Moscovitch, M., & Levine, B. (2012). Personal semantics: At the crossroads of semantic and episodic memory. *Trends in Cognitive Sciences*, *16*(11), 550–558. <https://doi.org/10.1016/j.tics.2012.09.003>
- Renoult, L., Irish, M., Moscovitch, M., & Rugg, M. D. (2019). From knowing to remembering: The semantic-episodic distinction. *Trends in Cognitive Sciences*, *23*(12), 1041–1057.
- Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *eLife*, *5*, e18260.
- Roberts, P., McGinnis, D., & Bladt, L. (1994). The unexpected excursion: Passive memories in everyday life. Poster presented at the 3rd Practical Aspects of Memory Conference, University of Maryland.
- Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented on lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814.
- Rubin, D. C. (1995). *Memory in oral traditions. The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford University Press.
- Rubin, D. C. (2022). A conceptual space for episodic and semantic memory. *Memory & Cognition*, *50*(3), 464–477. <https://doi.org/10.3758/s13421-021-01148-3>
- Rubin, D. C., & Berntsen, D. (2009). The frequency of voluntary and involuntary autobiographical memories across the life span. *Memory & Cognition*, *37*(5), 679–688.
- Rubin, D. C., & Umanath, S. (2015). Event memory: A theory of memory for laboratory, autobiographical, and fictional events. *Psychological Review*, *122*(1), 1–23. <https://doi.org/10.1037/a0037907>
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 251–257.
- Ryals, A. J., & Cleary, A. M. (2012). The recognition without cued recall phenomenon: Support for a feature-matching theory over a partial recollection account. *Journal of Memory and Language*, *66*, 747–762.
- Sanson, M., Cardwell, B. A., Rasmussen, A. S., & Garry, M. (2020). Evidence that “voluntary” versus “involuntary” retrieval is a fluency-based attribution. *Psychological Reports*, *123*(1), 141–158. <https://doi.org/10.1177/0033294119854180>

- Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: Remembering, imagining, and the brain. *Neuron*, 76(4), 677–694.
- Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory and Cognition*, 36, 920–932. doi:10.3758/MC.36.5.920
- Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, 15(7), 319–326.
- Schwartz, B. L. (2001). The relationship of tip-of-the-tongue states and retrieval time. *Memory & Cognition*, 29, 117–126.
- Schwartz, B. L. (2002). The strategic control of retrieval during tip-of-the-tongue states. *The Korean Journal of Thinking & Problem Solving*, 12, 27–37.
- Schwartz, B. L. (2011). The effect of being in a tip-of-the-tongue state on subsequent items. *Memory & Cognition*, 39, 245–250.
- Schwikert, S. R., & Curran, T. (2014). Familiarity and recollection in heuristic decision making. *Journal of Experimental Psychology: General*, 143(6), 2341.
- Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology*, 18(7), 791–807.
- Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters*, 680, 39–53.
- Simons, J. S., Ritchey, M., & Fernyhough, C. (2022). Brain mechanisms underlying the subjective experience of remembering. *Annual Review of Psychology*, 73, 159–186. https://doi.org/10.1146/annurev-psych-030221-025439
- Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, 66, 487–518. https://doi.org/10.1146/annurev-psych-010814-015331
- Smallwood, J., Tipper, C., Brown, K., Baird, B., Engen, H., Michaels, J. R., & Schooler, J. W. (2013). Escaping the here and now: Evidence for a role of the default mode network in perceptually decoupled thought. *Neuroimage*, 69, 120–125.
- Sonne, T., Kingo, O. S., Berntsen, D., & Krøjagaard, P. (2019). Thirty-five-month-old children have spontaneous memories despite change of context for retrieval. *Memory*, 27(1), 38–48. https://doi.org/10.1080/09658211.2017.1363243
- Souchay, C., & Moulin, C. J. A. (2009). Memory and consciousness in Alzheimer's disease. *Current Alzheimer Research*, 6, 186–195.
- Souchay, C., Moulin, C. J. A., Clarys, D., Tacconat, L., & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: Evidence from feeling of knowing and recollection. *Consciousness and Cognition*, 16, 769–784.
- Stawarczyk, D., Majerus, S., Maquet, P., & D'Argembeau, A. (2011). Neural correlates of ongoing conscious experience: Both task-unrelatedness and stimulus-independence are related to default network activity. *PLoS ONE*, 6(2), e16997.
- Suzuki, W. A., & Amaral, D. G. (1994). Perirhinal and parahippocampal cortices of the macaque monkey: Cortical afferents. *Journal of Comparative Neurology*, 350(4), 497–533. doi:10.1002/cne.903500402
- Tekcan, A.İ., Yılmaz, E., Kaya Kızıloğlu, B., Karadöller, D. Z., Mutafoglu, M., & Aktan Erciyes, A. (2015). Retrieval and phenomenology of autobiographical memories in blind individuals. *Memory*, 23(3), 329–339.
- Tsivilis, D., Allan, K., Roberts, J., Williams, N., Downes, J. J., & El-Dereby, W. (2015). Old-new ERP effects and remote memories: The late parietal effect is absent as recollection fails whereas the early mid-frontal effect persists as familiarity is retained. *Frontiers in Human Neuroscience*, 9, 532. doi:10.3389/fnhum.2015.00532
- Turner, M. S., Shores, E. A., Breen, N., & Coltheart, M. (2017). Déjà vu for news events but not personal events: A dissociation between autobiographical and non-autobiographical episodic memory processing. *Cortex*, 87, 142–155.
- Urquhart, J., & O'Connor, A. R. (2014). The awareness of novelty for strangely familiar words: A laboratory analogue of the déjà vu expérience. *PeerJ*, 2, e666.
- Urquhart, J. A., Sivakumaran, M. H., Macfarlane, J. A., & O'Connor, A. R. (2021). fMRI evidence supporting the role of memory conflict in the déjà vu experience. *Memory*, 29(7), 921–932. doi:10.1080/09658211.2018.1524496
- Uzer, T., & Brown, N. R. (2017). The effect of cue content on retrieval from autobiographical memory. *Acta Psychologica*, 172, 84–91. https://doi.org/10.1016/j.actpsy.2016.11.012
- Uzer, T., Lee, P. J., & Brown, N. R. (2012). On the prevalence of directly retrieved autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(5), 1296–1308. doi:10.1037/a0028142
- Van Calster, L., D'Argembeau, A., Salmon, E., Peters, F., & Majerus, S. (2017). Fluctuations of attentional networks and default mode network during the resting state reflect variations in cognitive states: Evidence from a novel resting-state experience sampling method. *Journal of Cognitive Neuroscience*, 29(1), 95–113.
- Vannucci, M., Batool, I., Pelagatti, C., & Mazzoni, G. (2014). Modifying the frequency and characteristics of involuntary autobiographical memories. *PLoS ONE*, 9(4), 1–7. https://doi.org/10.1371/journal.pone.0089582
- Vannucci, M., Pelagatti, C., Hanczakowski, M., & Chiorri, C. (2019). Visual attentional load affects the frequency of involuntary autobiographical memories and their level of meta-awareness. *Memory & Cognition*, 47, 117–129. doi:10.3758/s13421-018-0854-0
- Vannucci, M., Pelagatti, C., Hanczakowski, M., Mazzoni, G., & Paccani, C. R. (2015). Why are we not flooded by involuntary autobiographical memories? Few cues are more effective than many. *Psychological Research Psychologische Forschung*, 79(6), 1077–1085. https://doi.org/10.1007/s00426-014-0632-y
- Vignal, J. P., Maillard, L., McGonigal, A., & Chauvel, P. (2007). The dreamy state: Hallucinations of autobiographic memory evoked by temporal lobe stimulations and seizures. *Brain*, 130, 88–99.
- Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia*, 46(7), 1787–1799. https://doi.org/10.1016/j.neuropsychologia.2008.01.004
- Whittlesea, B. W., & Williams, L. D. (2001). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 14–33.
- Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica*, 98, 141–165.
- Wilckens, K. A., Erickson, K. I., & Wheeler, M. E. (2012). Age-related decline in controlled retrieval: The role of the PFC and sleep. *Neural Plasticity*, 2012, 624795. https://doi.org/10.1155/2012/624795
- Winkler, J. D., Kanouse, D. E., & Ware, J. E. (1982). Controlling for acquiescence response set in scale development. *Journal of Applied Psychology*, 67, 555–561. 17.
- Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114(1), 152.
- Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, 117(4), 1025–1054. https://doi.org/10.1037/a0020874
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46, 441–517.
- Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and response speed. *Canadian Journal of Experimental Psychology*, 48, 516–534.
- Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin*, 133(5), 800.
- Zedelius, C. M., Protzko, J., & Schooler, J. W. (2021). Lay theories of the wandering mind: Control-related beliefs predict mind wandering rates in- and outside the Lab. *Personality and Social Psychology Bulletin*, 47(6), 921–938. https://doi.org/10.1177/0146167220949408

Open Peer Commentary

On pattern completion, cues and future-oriented cognition

Donna Rose Addis^{a,b,c} and Karl K. Szpunar^d

^aRotman Research Institute, Baycrest Health Sciences, Toronto, ON, Canada;

^bDepartment of Psychology, University of Toronto, Toronto, ON, Canada;

^cSchool of Psychology, The University of Auckland, Auckland, New Zealand and

^dDepartment of Psychology, Toronto Metropolitan University, Toronto, ON, Canada

draddis@research.baycrest.org

karl.szpunar@ryerson.ca

https://www.baycrest.org/Baycrest/Research-Innovation/People/Researchers/Scientists/Dr-Donna-Rose-Addis

https://www.torontomu.ca/psychology/about-us/our-people/faculty/karl-szpunar/

doi:10.1017/S0140525X23000250, e357

Abstract

Barzykowski and Moulin's view on involuntary autobiographical memory focuses on automatic activation of representations and inhibitory control mechanisms. We discuss how and when a known neural mechanism – pattern completion – may result in involuntary autobiographical memories, the types of cues that may elicit this phenomenon and consider interactions with future-oriented cognition.

Barzykowski and Moulin (B&M) synthesize considerable evidence to support their primary thesis that involuntary autobiographical memories (IAMs) occur in the context of constant and automatic activation of autobiographical fragments in the absence of inhibitory control mechanisms; when the activation of a given memory representation passes a threshold, it is involuntarily retrieved. While we agree with their proposed model, here we discuss how another critical mechanism in direct retrieval – pattern completion – may result in IAMs, the types of cues that may elicit this phenomenon and how this process interacts with future-related cognition.

Pattern completion, the process of recalling an entire memory representation when cued with a subset of its elements, was originally proposed by Marr (1971) as a fundamental computation of the hippocampus along with pattern separation. He argued that whether incoming information elicits pattern completion depends on the degree to which it matches a previously stored representation, suggesting that as few as one-third of the elements may be sufficient to elicit recall (Becker, 2016). We suggest that B&M's proposal is broadly consistent with this and later computational models of pattern completion and separation (McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003) in two ways. First, the hippocampus is argued to behave in a thresholded manner depending on the overlap of the cue with the stored memory: If the threshold is met, the cue is pattern completed and retrieval ensues (Elfmán, Aly, & Yonelinas, 2014; see Eichenbaum, Yonelinas, & Ranganath, 2007, for a review of neuroimaging evidence of thresholded hippocampal signals). Second, once this threshold is met, retrieval occurs *automatically* via an autoassociation network mediated by the dense recurrent connections within CA3, resulting in a pattern of excitation in the output layer which is then back-projected to the complete set of neocortical regions representing the memory (Rolls, 2016; for a recent review of empirical evidence, see Becker, 2016).

Therefore, we argue that IAMs occur when content received by the hippocampus overlaps sufficiently with an existing memory as to trigger the automatic “completion” of the whole representation in neocortical regions. While conceptually similar to the proposed threshold model, in that overlap between the cue and the content must pass a threshold, pattern completion reflects a known neural mechanism by which automatic retrieval can occur. As a fundamental retrieval mechanism, pattern completion should act at all levels of the memory hierarchy (Conway & Pleydell-Pearce, 2000), including not just specific events but also general events and lifetime periods shown to engage the hippocampus (e.g., Ford, Addis, & Giovanello, 2011).

Although B&M suggest that cues are central to their account of spontaneous memory phenomena, little consideration is given to why some cues, whether present in the external environment or generated internally by the individual, are more likely to evoke IAMs than other cues. It is likely that, in everyday situations, a specific combination of cues is required to provide sufficient overlap with the stored representation in order for pattern completion to occur (Marr, 1971), explaining why IAMs do not occur all the time. Moreover, it may be that the interaction between cues and internal states is important. Along these lines, Klinger and colleagues have suggested that current concerns, or the state of an organism between commitment to a goal and later accomplishment or abandonment of that goal (Klinger, 1978), make goal-relevant cues especially salient and likely to evoke spontaneous thoughts and memories (e.g., Klinger & Cox, 2004).

Consider an individual who takes the same train ride into work each day. While looking out the window of the train, the individual is exposed to many cues (e.g., a billboard featuring an airplane; a junkyard full of old automobiles) that could, but typically do not, produce spontaneous retrieval of past experiences (Berntsen, 2009). However, if the individual were to set a goal (e.g., a planned trip) related to one of those cues (e.g., a billboard featuring an airplane), then that cue may be more likely to evoke not only thoughts about the goal, but conceivably some relevant experience from the past (e.g., a previous trip). Of course, in this scenario, the memory that comes to be evoked by the now-relevant cue may be more likely to surpass some threshold of activation due to its relation to the goal, but it is important not to overlook the fact that the cue itself is also more salient (Klinger, 2013). Understanding how goals, cues, and representations independently and/or jointly determine the occurrence of spontaneous memory phenomena, such as IAMs, are issues warranting further attention in the literature (e.g., Cole & Berntsen, 2016).

Consideration of goals alongside cues and representations may help shed light on other related phenomena. For instance, one important function of memory is to retain simulations related to predictions, intentions, and plans for the future (Ingvar, 1985; Szpunar, Addis, McLelland, & Schacter, 2013). Recent studies have shown that simulations of the future are especially likely to come to mind in a spontaneous manner if they have been thought about on some previous occasion (e.g., Jeunehomme & D'Argembeau, 2016). This is likely the case because simulations result in memory representations that can be directly elicited by goal-relevant cues (for similar considerations in the prospective memory literature, see McDaniel & Einstein, 2000; for more general considerations of the role of pattern completion mechanisms in future-oriented cognition, see Falandays, Nguyen, & Spivey, 2021). Although more speculative, it is possible that in some situations, goal-relevant cues only weakly evoke memories of the future, along the lines of the continuum proposed by B&M, giving rise to a spontaneous familiarity-based feeling that one should be doing something in the present moment (a “*déjà vu* for the future”) as opposed to the more classic feeling that one has already experienced the present moment.

In sum, for a more complete account of IAMs and related phenomena, it will be important to consider more precisely the neural mechanism(s) governing the interaction between cues and representations, and *why* specific cues and representations are especially relevant and/or accessible at any moment in time.

Financial support. D. R. A. is funded by the Canada 150 Research Chairs Program. K. K. S. is funded by an Early Researcher Award from the Government of Ontario, the Natural Sciences and Engineering Research Council of Canada (RGPIN-2020-04618) and the Social Science and Humanities Research Council of Canada (430-2020-00515).

Competing interest. None.

References

- Becker, S. (2016). Marr's theory of the hippocampus as a simple memory: Decades of subsequent research suggest it is not that simple. In L. M. Vaina & R. E. Passingham (Eds.), *Computational theories and their implementation in the brain: The legacy of David Marr* (pp. 159–178). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198749783.003.0007>
- Berntsen, D. (2009). *Involuntary autobiographical memories: An introduction to the unbidden past*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511575921>

- Cole, S. N., & Berntsen, D. (2016). Do future thoughts reflect personal goals? Current concerns and mental time travel into the past and future. *Quarterly Journal of Experimental Psychology*, 69, 273–284. <https://doi.org/10.1080/17470218.2015.1044542>
- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107, 261–288. <https://doi.org/10.1037/0033-295X.107.2.261>
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 30, 123–152. <https://doi.org/10.1146/annurev.neuro.30.051606.094328>
- Elfman, K. W., Aly, M., & Yonelinas, A. P. (2014). Neurocomputational account of memory and perception: Thresholded and graded signals in the hippocampus. *Hippocampus*, 24, 1672–1686. <https://doi.org/10.1002/hipo.22345>
- Falandays, J. B., Nguyen, B. T., & Spivey, M. J. (2021). Is prediction nothing more than multi-scale pattern completion of the future? *Brain Research*, 1768, 147578. <https://doi.org/10.1016/j.brainres.2021.147578>
- Ford, J. H., Addis, D. R., & Giovanello, K. S. (2011). Differential neural activity during search of specific and general autobiographical memories elicited by musical cues. *Neuropsychologia*, 49, 2514–2526. <https://doi.org/10.1016/j.neuropsychologia.2011.04.032>
- Ingvar, D. H. (1985). “Memory of the future”: An essay on the temporal organization of conscious awareness. *Human Neurobiology*, 4, 127–136.
- Jeunehomme, O., & D’Argembeau, A. (2016). Prevalence and determinants of direct and generative modes of production of episodic future thoughts in the word cueing paradigm. *Quarterly Journal of Experimental Psychology*, 69, 254–272. <https://doi.org/10.1080/17470218.2014.993663>
- Klinger, E. (1978). Modes of normal conscious flow. In K. S. Pope & J. L. Singer (Eds.), *The stream of consciousness* (pp. 225–258). Plenum. https://doi.org/10.1007/978-1-4684-2466-9_9
- Klinger, E. (2013). Goal commitments and the content of thoughts and dreams: Basic principles. *Frontiers in Psychology*, 4, 415. <https://doi.org/10.3389/fpsyg.2013.00415>
- Klinger, E., & Cox, W. M. (2004). Motivation and the theory of current concerns. In W. M. Cox & E. Klinger (Eds.), *Handbook of motivational counseling: Concepts, approaches, and assessment* (pp. 3–27). John Wiley & Sons Ltd. <https://doi.org/10.1002/9780470713129.ch1>
- Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 262, 23–81. <https://doi.org/10.1098/rstb.1971.0078>
- McClelland, J. L., McNaughton, B. L., & O’Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419–457. <https://doi.org/10.1037/0033-295X.102.3.419>
- McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, 14, S127–S144. <https://doi.org/10.1002/acp.775>
- Norman, K. A., & O’Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, 110, 611–646. <https://doi.org/10.1037/0033-295X.110.4.611>
- Rolls, E. T. (2016). Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiology of Learning and Memory*, 129, 4–28. <https://doi.org/10.1016/j.nlm.2015.07.008>
- Szpunar, K. K., Addis, D. R., McLelland, V. C., & Schacter, D. L. (2013). Memories of the future: New insights into the adaptive value of episodic memory. *Frontiers in Behavioral Neuroscience*, 7, 47. <https://doi.org/10.3389/fnbeh.2013.0004>

Accounting for the strangeness, infrequency, and suddenness of déjà vu

Nikola Andonovski  and Kourken Michaelian 

Centre de philosophie de la mémoire, Institut de philosophie de Grenoble, Université Grenoble Alpes, CS 40700, Grenoble, Cedex 9, France
nikola.andonovski@univ-grenoble-alpes.fr
kourken.michaelian@univ-grenoble-alpes.fr
<http://phil-mem.org/members/andonovski.php>
<http://phil-mem.org/members/michaelian.php>

doi:10.1017/S0140525X23000237, e358

Abstract

Barzykowski and Moulin argue that déjà vu is a natural product of autobiographical memory retrieval. Their proposal fails to account for three salient properties of déjà vu experiences: Their strangeness, their infrequency, and their characteristically sudden onset. Accounting for these properties is necessary for proper integration of déjà vu into autobiographical memory research.

Aiming to provide a means of integrating research on the two phenomena, Barzykowski and Moulin (B&M) propose an approach to involuntary autobiographical memory and déjà vu as natural products of memory retrieval. Their proposal must, if it is to achieve this aim, have a level of detail sufficient to account for the core functional and phenomenological properties of the target phenomena. Focusing on déjà vu, this commentary shows that the proposal does not have the requisite level of detail. We argue, after identifying three core properties of déjà vu experiences, that each of these leads to a problem that the proposed account, in its current form, is unable to resolve.

The first property that we will consider is strangeness. Déjà vu experiences are regularly characterized as strange, weird, or even eerie by both subjects and theorists (e.g., Brázdil et al., 2012). In the kind of experience on which we will focus here – sometimes referred to as “déjà vécu” (O’Connor, Lever, & Moulin, 2010) – the strangeness of the experience is due to the fact that it concerns a singular event and not merely a repeatable item, “as if time had slipped a cog and were now repeating itself” (Woodworth, 1940: 357). The second property is infrequency. While most people experience déjà vu, they do so very rarely – on the order of few times a year (Brown, 2003). The final property is suddenness. Triggered by a variety of situational factors, déjà vu experiences typically begin abruptly, last no longer than a few seconds, and end just as abruptly as they began (Brown, Porter, & Nix, 1994).

Consider *strangeness* first. According to B&M’s proposal, the distinctive phenomenology of déjà vu results from the combination of incomplete memory retrieval, which produces a sense of familiarity, with metacognitive appraisal of that familiarity as implausible. The strangeness of déjà vu, however, is due not merely to the *implausibility* of familiarity with a repeatable item – that is, with a type – but also, in cases in which déjà vu amounts to déjà vécu, to the felt *impossibility* of familiarity with the currently experienced token event. In order to account for the strangeness of déjà vu, then, the proposal needs to explain not only why incomplete retrieval produces a sense of familiarity with an item (a feeling that the current event resembles something experienced in the past) but also why, at least in some paradigmatic cases of déjà vu, unexpected familiarity is not simply brushed off as arising from limited access but is instead taken to indicate event repetition (a feeling that this very event has already been experienced). If the distinctive phenomenology of déjà vu results from a metacognitive assessment of plausibility, in short, then what should be assessed is not the plausibility of event resemblance but of event repetition.

For a similar reason, B&M’s proposal cannot account for the *infrequency* of déjà vu. According to the proposal, incomplete retrieval – producing familiarity but not specific memory content – should be relatively frequent. This is required to explain the relative frequency of tip-of-the-tongue experiences (Brown, 1991).

Now, if our argument above is on the right track, the metacognitive assessment of plausibility responsible for generating déjà vu pertains to event repetition, not to mere event resemblance. Since event repetition is impossible, however, familiarity unaccompanied by specific memory content should typically, if not invariably, be assessed as impossible or at least highly implausible. But this entails that déjà vu experiences should be relatively frequent, which strongly contradicts the available data. Thus, if it is to account for both the strangeness and the infrequency of déjà vu, the proposal will have to invoke a wider range of (meta)cognitive processes. We will not attempt to determine here with it is feasible for a modified form of the proposal to do so.

The *suddenness* of déjà vu also presents a problem for the proposal. B&M contrast the rapidity of content retrieval in involuntary memory with the gradual intensification of familiarity in déjà vu. The onset of familiarity is sudden, but the epistemic feeling characteristic of déjà vu is a result of the conflict between familiarity and metacognitive expectation, which triggers additional search. Hence, as the authors suggest, the feeling of unexpected familiarity in déjà vu should gradually intensify until it is resolved or explained. (Compare to the gradual intensification, and persistence, of familiarity in tip-of-the-tongue experiences.) But déjà vu experiences have a characteristically sudden onset: Not only do they begin abruptly, but often the strangeness that characterizes them is strongest in the first moments of the experience. The proposal, at least in its current form, does not account for this property of déjà vu.

Before concluding, we note that the strangeness, infrequency, and suddenness of déjà vu have sometimes been viewed as pointing to an underlying neural or cognitive malfunction (e.g., Critchley, 1989). Their characterization of déjà vu as a *natural* product of memory retrieval suggests that B&M intend to argue against such views. If, on the one hand, their intention is to argue that déjà vu results from properly functioning retrieval processes, much more evidence is required to support their claim. If, on the other hand, they do not mean to argue for a claim about function, then their characterization of déjà vu as “natural” adds little to their proposal. Greater clarity about this issue would be welcome.

B&M might respond by arguing that our focus on déjà vu for token events (i.e., on déjà vécu) is unjustified, suggesting that future work, and the development of more detailed models, will shed more light on this particular phenomenon. While we sympathize with this sentiment, we worry that operationalizing déjà vu as any form of inappropriate familiarity (cf. Neppe, 1983) obscures much of what is puzzling, and indeed difficult to explain, about the phenomenon. We thus maintain that integrating déjà vu into autobiographical memory research will require tackling the problems that we have highlighted head-on.

Acknowledgements. The authors thank Denis Perrin for helpful discussion.

Financial support. N. A.’s research is funded by the European Union’s Horizon research and innovation program under the Marie Skłodowska-Curie grant agreement 101062754. K. M.’s research is supported by the Institut universitaire de France.


Competing interest. None.

References

- Brázdil, M., Mareček, R., Urbánek, T., Kašpárek, T., Mikl, M., Rektor, I., & Zeman, A. (2012). Unveiling the mystery of déjà vu: The structural anatomy of déjà vu. *Cortex*, 48(9), 1240–1243.
- Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109(2), 204–223.

- Brown, A. S. (2003). A review of the déjà vu experience. *Psychological Bulletin*, 129(3), 394–413.
- Brown, A. S., Porter, C. L., & Nix, L. A. (1994). *A questionnaire evaluation of the déjà vu experience*. Midwestern Psychological Association Convention.
- Critchley, E. M. R. (1989). The neurology of familiarity. *Behavioural Neurology*, 2(3), 195–200.
- Neppe, V. M. (1983). *The psychology of déjà vu: Have I been here before?* Witwatersrand University Press.
- O’Connor, A. R., Lever, C., & Moulin, C. J. (2010). Novel insights into false recollection: A model of déjà vécu. *Cognitive Neuropsychiatry*, 15(1–3), 118–144.
- Woodworth, R. S. (1940). *Psychology*. Henry Holt.

Neuropsychological predictions on involuntary autobiographical memory and déjà vu

Christine Bastin 

GIGA-Cyclotron Research Center-In Vivo Imaging (GIGA Institute), University of Liège, Liège, Belgium

Christine.Bastin@uliege.be
<https://www.gigaaml.uliege.be>

doi:10.1017/S0140525X23000249, e359

Abstract

I strongly support Barzykowski and Moulin in their proposal that common retrieval mechanisms can lead to distinct phenomenological memory experiences. I emphasize the importance of one of these mechanisms, namely the attribution system. Neuropsychological studies should help clarifying the role of these retrieval mechanisms, notably in cases of medial temporal-lobe lesions and cases of dementia.

A fundamental tenet in Barzykowski and Moulin’s (B&M) proposal is the idea of common retrieval mechanisms leading to distinct phenomenological memory experiences depending on the success/failure and degree of involvement of these mechanisms. In both involuntary autobiographical memory and déjà vu, there is a rapid cue-generated automatic search in memory, a feeling of familiarity and attribution processes that come into play. I agree with B&M that a key difference lies in the content reactivation, with content that comes successfully to mind in the case of involuntary autobiographical memories and failure to find any content in the case of déjà vu. However, more emphasis could be made on the fact that the two types of memory experiences critically differ in the extent of contribution of the attribution system. Of note, the notion of attribution system refers here to the cognitive appraisal of the result of the memory search to generate an output (Bastin et al., 2019). In other words, this corresponds to the set of inferential and monitoring processes that evaluate retrieved contents before deciding about the old/new status of the information or expressing a subjective feeling of memory. In involuntary autobiographical memories, the retrieval cue leads to the rapid reactivation of some content that is compatible with expectations that there is indeed a memory trace. Here, the match between content reactivation and feelings of familiarity is satisfactory and not surprising, so that attribution processes take the form of relatively automatic and unconscious inferential processes. In contrast, in déjà vu, some cues generate expectation of

processing fluency and feeling of familiarity, but the search fails to find any content. Attribution processes become hyperactive because they detect a mismatch between content and feeling. Moreover, there is knowledge that it is not possible to have content. This conflict reaches consciousness and is likely responsible for the unique phenomenology of déjà vu. Another common mechanism that is evoked in B&M is that elements of internal and external context are automatically weighted against prior knowledge, and this generates expectations. There are similarities with the ideas proposed in the predictive interactive multiple memory systems model (Henson & Gagnepain, 2010). Notably, Henson and Gagnepain (2010) suggest that the content retrieved in a memory system is compared to context-based predictions. In this view, déjà vu arises from a prediction error as the absence of memory content mismatches the expectations.

If involuntary autobiographical memory and déjà vu rely on the same neurocognitive architecture, one expects that selective dysfunction of some retrieval mechanisms following brain damage impacts the expression of these phenomenological experiences. Predictions for brain-damaged patients is only briefly mentioned in B&M's article, but this is an important avenue for future research as it provides a way to test predictions emerging from the proposal. B&M predict that patients with medial temporal-lobe damage should show similar changes in déjà vu and involuntary autobiographical memories. For instance, temporal-lobe epilepsy would be associated with more déjà vu experiences and more involuntary autobiographical memories. However, if the two memory phenomena differ critically on the access to content, one may predict a different pattern. Considering that the medial temporal lobe is critical to store contents of past episodes (Davachi, 2006; Eichenbaum, Yonelinas, & Ranganath, 2007; Ranganath & Ritchey, 2012), medial temporal-lobe damage should decrease involuntary autobiographical memories; and as metacognitive appraisal would partly depend on frontal areas (Chua, Pergolizzi, & Weintraub, 2014; Kurilla & Gonsalves, 2012) and is not expected to be impaired in medial temporal-lobe pathology, such patients may show preserved or more déjà vu experiences. The opposite pattern would be expected in patients with frontal-lobe damage, with the disappearance of déjà vu experiences and intact or increased involuntary autobiographical memories.

An interesting case is the one of patients with dementia as the extent of the neuropathology disrupts several retrieval mechanisms evoked as the building blocks for involuntary autobiographical memory and déjà vu. Patients with dementia sometimes present with persistent déjà vu whereby they have the feeling that everyday life events constantly repeat (Moulin, 2013). These patients typically come up with some explanations about their pervasive feeling of memory (i.e., confabulations) and do not report knowing that this feeling of memory is false. This is likely because both memory content and attribution processes are altered because of medial temporal and frontal pathology (Moulin, 2013). Another line of research suggests that demented patients have preserved involuntary autobiographical memories evoked by nostalgia films or music (El Haj, Fasotti, & Allain, 2012; Rasmussen, Salgado, Daustrand, & Berntsen, 2021), which may seem surprising if these memories rely on the same retrieval mechanisms that are impaired and cause persistent déjà vu. Nevertheless, it is likely that those memories are recollection of remote memories, mostly belonging to young adulthood, whose content is consolidated (Berntsen, Kirk, & Kopelman, 2022). So, the question remains as to the integrity of involuntary autobiographical recall of recent memories in demented patients.

This also raises the possibility that involuntary autobiographical memory and déjà vu differ on another dimension, which is temporality. Whereas involuntary autobiographical memory brings back to mind events from any time in the past, déjà vu seems anchored in the present (i.e., “the situation I experience now feels familiar even if I know it is new”) with some glimpse into the future with the feeling of prescience. I agree with B&M that considering both involuntary autobiographical memory and déjà vu experiences within the same sample would provide insights as to their common mechanisms.

Financial support. This work was supported by the F.R.S.-FNRS (C. B. is a senior research associate at F.R.S.-FNRS).

Competing interest. None.

References

- Bastin, C., Besson, G., Simon, J., Delhay, E., Geurten, M., Willems, S., & Salmon, E. (2019). An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, 42(e281), 1–60. <https://doi.org/10.1017/S0140525X19000621>
- Berntsen, D., Kirk, M., & Kopelman, M. D. (2022). Autobiographical memory loss in Alzheimer's disease: The role of the reminiscence bump. *Cortex*, 150, 137–148. <https://doi.org/10.1016/j.cortex.2022.02.008>
- Chua, E., Pergolizzi, D., & Weintraub, R. R. (2014). The cognitive neuroscience of meta-memory monitoring: Understanding metamemory processes, subjective levels expressed, and metacognitive accuracy. In S. M. Fleming & C. D. Frith (Eds.), *The cognitive neuroscience of metacognition* (pp. 267–291). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-45190-4_12
- Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, 16(6), 693–700. <https://doi.org/10.1016/j.conb.2006.10.012>
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 30, 123–152. <https://doi.org/10.1146/annurev.neuro.30.051606.094328>
- El Haj, M., Fasotti, L., & Allain, P. (2012). The involuntary nature of music-evoked autobiographical memories in Alzheimer's disease. *Consciousness and Cognition*, 21(1), 238–246. <https://doi.org/10.1016/j.concog.2011.12.005>
- Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, 20(11), 1315–1326. <https://doi.org/10.1002/hipo.20857>
- Kurilla, B. P., & Gonsalves, B. D. (2012). An ERP investigation into the strategic regulation of the fluency heuristic during recognition memory [Research Support, N.I.H., Extramural]. *Brain Research*, 1442, 36–46. <https://doi.org/10.1016/j.brainres.2011.12.060>
- Moulin, C. J. (2013). Disordered recognition memory: Recollective confabulation. *Cortex*, 49(6), 1541–1552. <https://doi.org/10.1016/j.cortex.2013.01.010>
- Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews: Neuroscience*, 13, 1–15.
- Rasmussen, K. W., Salgado, S., Daustrand, M., & Berntsen, D. (2021). Using nostalgia films to stimulate spontaneous autobiographical remembering in Alzheimer's disease. *Journal of Applied Research in Memory and Cognition*, 10(3), 400–411. <https://doi.org/10.1016/j.jarmac.2020.11.001>

Does inhibitory (dys)function account for involuntary autobiographical memory and déjà vu experience?

Thomas F. Burns 

Okinawa Institute of Science and Technology Graduate University, Onna-son, Kunigami-gun, Okinawa, Japan
t.f.burns@gmail.com
<https://tfburns.com/>

doi:10.1017/S0140525X23000146, e360

Abstract

External cues and internal configuration states are the likely instigators of involuntary autobiographical memories (IAMs) and déjà vu experience. Indeed, Barzykowski and Moulin discuss relevant neuroscientific evidence in this direction. A complementary line of enquiry and evidence is the study of inhibition and its role in memory retrieval, and particularly how its (dys)function may contribute to IAMs and déjà vu.

Barzykowski and Moulin (B&M) suggest that involuntary autobiographical memories (IAMs) and the experience of déjà vu can be described and explained as natural phenomena resulting from memory processing, particularly retrieval. Both are based on the same basic retrieval processes, they argue, and can be considered as “involuntary” or spontaneous cognitive events. The target paper discusses similarities and differences between the two phenomena, and outlines how these unusual subjective experiences may be explained by relating them to known mechanisms of memory retrieval. This leads the authors to conclude that IAMs and déjà vu lie on a continuum. Here, I expand on why inhibitory (dys)function may play a significant role in such a continuum.

In memory systems, distinct patterns of neuronal activity (Atallah & Scanziani, 2009; Colgin, 2016) are modulated intricately by inhibitory neurons in the hippocampus and cortex. More generally, inhibitory neurons are important for the structural formation and proper functioning of neuronal assemblies (Holtmaat & Caroni, 2016). This is due to their fast-spiking behaviour which allows them to provide powerful feedforward and feedback inhibition to excitatory neurons (Gan, Ming Weng, Pernía-Andrade, Csicsvari, & Jonas, 2017) and constrain the size or identity of such assemblies (Colgin, 2016; Gan et al., 2017; Isaacson & Scanziani, 2011). This limits the realm of excitatory plasticity for neurons in such assemblies, subsequently providing increased efficiency for memory storage and greater robustness against over-excitation (Mongillo, Rumpel, & Loewenstein, 2018), which may also help reduce the number of spurious attractors. From a computational perspective, it therefore seems natural that inhibition will be implicated in IAMs and déjà vu.

From a psychological perspective, Conway and Pleydell-Pearce (2000) describe a model wherein a cognitive inhibitory control mechanism may actively suppress the activation of memories while they are not relevant. Such a mechanism could prevent such memories from becoming fully activated in our conscious awareness, while leaving room for the experience of familiarity in the absence of the underlying memory content. Retrieval of IAMs could occur via a similar mechanism. While B&M do not categorically reject this as an explanation, they point to recent studies (Barzykowski, Radel, Niedźwieńska, & Kvavilashvili, 2019; Barzykowski, Staugaard, & Mazzoni, 2021) which suggest a retrieval threshold or activation account. For example, IAMs may be particularly effective in capturing memory-related attention due, for example, to certain phenomenological properties such as emotional valence or particularity. However, it seems more likely that both accounts are true and work in concert.

In computational associative memory models where memories are spatially correlated (Burns, Haga, & Fukai, 2022; Haga & Fukai, 2019), memory capacity and the range of retrieval across related memories is dramatically altered by subtle modulations

of inhibition – without any change to the underlying excitation or memory structures. While there is still a critical threshold or level of activation which must be obtained for memories to become activated, the dynamics of these mechanisms are changed by inhibitory modulation. In the context of humans, where inhibition protects against memory interference (Koolschijn et al., 2019), this may explain why clinical studies have reported that delusions and hallucinations associated with schizophrenia are due to impaired inhibitory regulation (Vogels & Abbott, 2007; Yizhar & Fenno, 2011).

Learning of context-dependent associations has also been shown in rats to be correlated with changes in inhibition (Kuchibhotla et al., 2017) and inhibition may play an active role in helping to forget competing memories of a retrieved memory (Wimber, Alink, Charest, Kriegeskorte, & Anderson, 2015), perhaps *pruning* the relevant assemblies to improve or maintain the strength of the retrieved memory. It seems likely that IAMs and déjà vu may be implicated or are affected by this “pruning” role, in combination with other inhibitory mechanisms, for example, disinhibition (Letzkus, Wolff, & Lüthi, 2015), and that these phenomena will subsequently occur in different frequencies for persons with inhibitory neuron disorders. Inhibitory function is also known to change when exposed to different drugs, and indeed we see that recreational uses of alcohol and other substances often increase mind wandering (Sayette, 2009), a phenomenon the authors argue is related to IAMs and déjà vu.

In summary, the role of inhibitory neurons in memory systems is well-established. These neurons are important for the structural formation and proper functioning of neuronal assemblies, which allows for efficient memory storage and increased robustness against over-excitation. Inhibition also plays a role in protecting against memory interference and competition. Dysfunction in inhibitory regulation has been linked to delusions and hallucinations, and subtle modulation of inhibition shows dramatic changes in memory retrieval dynamics in computational models. These findings suggest that IAMs and déjà vu experiences may be related to the function of inhibitory neurons or their cognitive control. These phenomena may occur at different frequencies for individuals with disorders affecting inhibitory function.

Unifying the threshold/activation and inhibitory control accounts of IAMs and déjà vu thus appears a sensible goal; it is encouraging the authors state that “our position is relatively inclusive” in this regard and do not categorically reject the inhibitory account. The challenge will be to design suitable experiments with relevant human populations and in relevant conditions to connect these high-level subjective phenomena with the rapidly growing biological and computational literature non-trivially implicating inhibition in memory processes.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interest. None.

References

- Atallah, B. V., & Scanziani, M. (2009). Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition. *Neuron*, 62, 566–577.
- Barzykowski, K., Radel, R., Niedźwieńska, A., & Kvavilashvili, L. (2019). Why are we not flooded by involuntary thoughts about the past and future? Testing the cognitive inhibition dependency hypothesis. *Psychological Research*, 83(4), 666–683.
- Barzykowski, K., Staugaard, S. R., & Mazzoni, G. (2021). Retrieval effort or intention: Which is more important for participants’ classification of involuntary and voluntary memories? *British Journal of Psychology*, 112(4), 1080–1102.

- Burns, T. F., Haga, T., & Fukai, T. (2022). Multiscale and extended retrieval of associative memory structures in a cortical model of local-global inhibition balance. *eNeuro*, 9(3), ENEURO.0023-22.2022. <https://doi.org/10.1523/ENEURO.0023-22.2022>
- Colgin, L. L. (2016). Rhythms of the hippocampal network. *Nature Reviews Neuroscience*, 17, 239–249.
- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107(2), 261.
- Gan, J., Ming Weng, S., Pernía-Andrade, A. J., Csicsvari, J., & Jonas, P. (2017). Phase-locked inhibition, but not excitation, underlies hippocampal ripple oscillations in awake mice in vivo. *Neuron*, 93, 308–314.
- Haga, T., & Fukai, T. (2019). Extended temporal association memory by modulations of inhibitory circuits. *Physical Review Letters*, 123, 078101.
- Holtmaat, A., & Caroni, P. (2016). Functional and structural underpinnings of neuronal assembly formation in learning. *Nature Neuroscience*, 19, 1553–1562.
- Isaacson, J. S., & Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron*, 72, 231–243.
- Koolschijn, R. S., Emir, U. E., Pantelides, A. C., Nili, H., Behrens, T. E., & Barron, H. C. (2019). The hippocampus and neocortical inhibitory engrams protect against memory interference. *Neuron*, 101, 528–541.e6.
- Kuchibhotla, K. V., Gill, J. V., Lindsay, G. W., Papadoyannis, E. S., Field, R. E., Sten, T. A., ... Froemke, R. C. (2017). Parallel processing by cortical inhibition enables context-dependent behavior. *Nature Neuroscience*, 20, 62–71.
- Letzkus, J. J., Wolff, S. B., & Lüthi, A. (2015). Disinhibition, a circuit mechanism for associative learning and memory. *Neuron*, 88, 264–276.
- Mongillo, G., Rumpel, S., & Loewenstein, Y. (2018). Inhibitory connectivity defines the realm of excitatory plasticity. *Nature Neuroscience*, 21, 1463–1470.
- Sayette, M. A., Reichle, E. D., & Schooler, J. W. (2009). Lost in the sauce: The effects of alcohol on mind wandering. *Psychological Science*, 20(6), 747–752. <https://doi.org/10.1111/j.1467-9280.2009.02351.x>
- Vogels, T., & Abbott, L. (2007). Gating deficits in model networks: A path to schizophrenia? *Pharmacopsychiatry*, 40, S73.
- Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., & Anderson, M. C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature Neuroscience*, 18, 582–589.
- Yizhar, O., & Fenno, L. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, 477, 171–178.

A possible shared underlying mechanism among involuntary autobiographical memory and déjà vu

Anne M. Cleary^a , Cati Poulos^b and Caitlin Mills^b

^aDepartment of Psychology, Colorado State University, Fort Collins, CO, USA and ^bDepartment of Educational Psychology, University of Minnesota, Minneapolis, MN, USA

Anne.Cleary@colostate.edu

mcp1066@wildcats.unh.edu

caitlin.s.mills@gmail.com

[https://psywebserv.psych.colostate.edu/Psylist/facdetail.php?](https://psywebserv.psych.colostate.edu/Psylist/facdetail.php?FirstName=Anne&LastName=Cleary)

[FirstName=Anne&LastName=Cleary](https://www.cehd.umn.edu/edpsych/people/cmills/)

<https://www.cehd.umn.edu/edpsych/people/cmills/>

doi:10.1017/S0140525X23000079, e361

Abstract

We propose that IAM and déjà vu may not share a placement on the same gradient, per se, but the mechanism of *cue familiarity detection*, and a major differentiating factor between the two metacognitive experiences is whether the resulting inward directed search of memory yields retrieved content or not. Déjà vu may manifest when contentless familiarity detection is inexplicable by the experimenter.

In their article, “Are involuntary autobiographical memory and déjà vu natural products of memory retrieval?” Barzykowski and Moulin (B&M) bring two fascinating and usually disparate fields together in a way that we believe will be fruitful for future research and theory. The idea that déjà vu should be considered a form of involuntary memory makes great sense given that déjà vu interrupts ongoing processing and pulls attention toward the experience of déjà vu itself; also, both déjà vu and involuntary autobiographical memory (IAM) are arguably phenomenologically surprising in nature (Christoff, Irving, Fox, Spreng, & Andrews-Hanna, 2016; Mills, Zamani, White, & Christoff, 2021).

Although the linkage made in the article is already exciting, a potentially helpful extension for the field will be to specify the common mechanism(s) that might be shared between IAM and déjà vu, as well as what distinguishes them phenomenologically (e.g., Neisser et al., 2023). We propose that IAM and déjà vu may not share a placement on the same gradient, per se, but rather the mechanism of *cue familiarity detection*. Upon detecting familiarity with a cue, an inward directed attentional search of memory may be launched that can manifest in the form of attempts at recollecting potentially relevant candidate information (Carlaw, Huebert, McNeely-White, Rhodes, & Cleary, 2022; Huebert, McNeely-White, & Cleary, 2023). We propose that cue familiarity detection can be triggered involuntarily in the presence of a cue or set of cues, which can be internally or externally driven. The familiarity “signal” that emerges varies according to the degree of feature overlap between the cue(s) and memory representations (Clark & Gronlund, 1996; McNeely-White, McNeely-White, & Cleary, 2021; McNeely-White, McNeely-White, Huebert, Carlaw, & Cleary, 2022); if it is above a critical threshold, attention may be directed inward toward a search of memory to attempt to retrieve candidate relevant information to the current situation.

This can take at least two different forms, which are similar as noted by B&M. Sometimes, the search launched by initial cue familiarity detection may result in successful retrieval of a relevant prior experience from memory, leading to an experience of an IAM. Notably, although most IAMs have identifiable cues (Ball & Little, 2006; Schlagman & Kvavilashvili, 2008), it is also possible that the cue cannot be identified (the cue-elicited content simply comes to mind in response to the cue, even if the cue itself was not identified by the experimenter). Other times, the search that is launched by initial cue familiarity detection fails. When it fails, we propose that attention is drawn to the sense of familiarity that prompted the inward directed attentional search of memory. In the absence of any retrieved memory content, the experience might be that of either a general familiarity feeling (if the familiarity seems like it is explainable) or déjà vu (if the familiarity cannot easily be explained away). From this perspective, one possible common memory-based mechanism between IAM and déjà vu is cue familiarity detection, and a major differentiating factor between the two is whether the inward directed search of memory results in retrieved content or not.

With this in mind, one aspect that may phenomenologically distinguish IAM and déjà vu is the presence versus the absence of content (see Neisser et al., 2023). In the case of IAM, content is retrieved that can usually be attributed to a cue or cues that elicited it roughly 85% of the time (Ball & Little, 2006; Schlagman & Kvavilashvili, 2008). In the case of déjà vu, no content is retrieved, leaving a person in a state of contentless inward focus and a mere sensation of a memory; this sensation can be difficult to explain away when there is no content identified

and it is unclear what elicited it, which can be off-putting. The inability to explain the feeling may contribute to a more salient phenomenological experience of surprise or spontaneity in the thought stream (Mills et al., 2021).

In this regard, déjà vu may fit within Whittlesea and Williams's (2000, 2001a, 2001b) discrepancy attribution framework in that it may be when the familiarity feeling is surprising that déjà vu occurs; however, as B&M note, no laboratory paradigm has yet successfully operationalized "surprise" in this context, and existing attempts at doing so have been unsuccessful (see Cleary, Morris, & Langley, 2007; Karpicke, McCabe, & Roediger, 2008). We propose that it is the inexplicable contentless sensation of memory that characterizes déjà vu. When people experience a strong sensation of familiarity, they have an inherent need to explain why the feeling is occurring, and may do so by conjuring up details that are inaccurate (Carlaw et al., 2022; Huebert et al., 2023) or even confabulating details (Moulin, 2013, 2018) in an effort to explain away the feeling. The more difficult it is to do this, the more likely the experience might be labeled as déjà vu as opposed to just a feeling of familiarity.

Some evidence for this possible mechanistic link can be found in other shared memory-based characteristics of IAM and déjà vu. For example, a shared characteristic between IAM and déjà vu is that both are associated with medial temporal lobe (MTL) processes (Christoff et al., 2016; Cleary & Brown, 2022). It is plausible that this commonality is due to MTL processes being responsible for the initial familiarity detection that launches the search of memory; some researchers have indeed proposed that familiarity detection results from processing occurring in the MTL (e.g., Elfman & Yonelinas, 2015). From a phenomenological perspective, both IAM and déjà vu are more commonly reported in younger adults compared to older adults (e.g., Cleary & Brown, 2022; Schlagman, Kvavilashvili, & Schulz, 2007). It is plausible that this commonality is due to the common mechanism of cue familiarity detection prompting initial memory search, and that this mechanism peaks in young adulthood and then diminishes with age.

In sum, we were delighted to see the interesting links made by B&M across two interesting involuntary and spontaneous thought processes. We think this work will inspire new theoretical and empirical work, and it has certainly prompted us to think more about their shared links, as discussed in this commentary.

References

Ball, C. T., & Little, J. C. (2006). A comparison of involuntary autobiographical memory retrievals. *Applied Cognitive Psychology, 20*(9), 1167–1179.

Carlaw, B. N., Huebert, A. M., McNeely-White, K. L., Rhodes, M. G., & Cleary, A. M. (2022). Detecting a familiar person behind the surgical mask: Recognition without identification among masked versus sunglasses-covered faces. *Cognitive Research: Principles and Implications, 7*(1), 90. <https://doi.org/10.1186/s41235-022-00440-3>

Christoff, K., Irving, Z. C., Fox, K. C., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience, 17*(11), 718–731. <https://doi.org/10.1038/nrn.2016.113>

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3*(1), 37–60.

Cleary, A. M., & Brown, A. S. (2022). *The déjà vu experience* (2nd ed.). Routledge.

Cleary, A. M., Morris, A. L., & Langley, M. M. (2007). Recognition memory for novel stimuli: The structural regularity hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 379–393.

Elfman, K. W., & Yonelinas, A. P. (2015). Recollection and familiarity exhibit dissociable similarity gradients: A test of the complementary learning systems model. *Journal of Cognitive Neuroscience, 27*(5), 876–892. https://doi.org/10.1162/jocn_a_00763

Huebert, A. M., McNeely-White, K. L., & Cleary, A. M. (2023). On the relationship between tip-of-the-tongue states and partial recollective experience: Illusory partial

recollective access during tip-of-the-tongue states. *Journal of Experimental Psychology: General, 152*(2), 542–570. <https://doi.org/10.1037/xge0001292>

Karpicke, J. D., McCabe, D. P., & Roediger, H. L. (2008). False memories are not surprising: The subjective experience of an associative memory illusion. *Journal of Memory and Language, 58*, 1065–1079.

McNeely-White, K. L., McNeely-White, D. G., & Cleary, A. M. (2021). Global matching in music familiarity: How musical features combine across memory traces to increase familiarity with the whole in which they are embedded. *Journal of Memory and Language, 118*, 104217. <https://doi.org/10.1016/j.jml.2020.104217>

McNeely-White, K. L., McNeely-White, D. G., Huebert, A. M., Carlaw, B. N., & Cleary, A. M. (2022). Specifying a relationship between semantic and episodic memory in the computation of a feature-based familiarity signal using MINERVA 2. *Memory & Cognition, 50*(3), 527–545. <https://doi.org/10.3758/s13421-021-01234-6>

Mills, C., Zamani, A., White, R. A., & Christoff, K. (2021). Out of the blue: Understanding abrupt and wayward transitions in thought through probability and predictive processing. *Philosophical Transactions of the Royal Society B, 376*(1817), 20190692. <https://doi.org/10.1098/rstb.2019.0692>

Moulin, C. (2018). *The cognitive neuropsychology of déjà vu*. Routledge.

Moulin, C. J. (2013). Disordered recognition memory: Recollective confabulation. *Cortex, 49*(6), 1541–1552.

Neisser, J., Abreu, G., Drane, D. L., Pedersen, N. P., Parsons, T. D., McNeely-White, K. L., & Cleary, A. M. (2023). Opening a conceptual space for metamemory experience. *New Ideas in Psychology, 69*, 100995. <https://doi.org/10.1016/j.newideapsych.2022.100995>

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition, 36*(5), 920–932.


Schlagman, S., Kvavilashvili, L., & Schulz, J. (2007). Effects of age on involuntary autobiographical memories. In J. H. Mace (Ed.), *Involuntary memory* (pp. 87–112). Blackwell Publishing. <https://doi.org/10.1002/9780470774069.ch5>

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 547–565.

Whittlesea, B. W. A., & Williams, L. D. (2001a). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings and familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 3–13.

Whittlesea, B. W. A., & Williams, L. D. (2001b). The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 14–33.

Intracranial electrical brain stimulation as an approach to studying the (dis)continuum of memory experiential phenomena

Jonathan Curot^{a,b,c} , Anaïs Servais^{a,c} and Emmanuel J. Barbeau^{a,c}

^aBrain and Cognition Research Center, CerCo, CNRS, UMR 5549, Toulouse, France; ^bDepartment of Neurology, Toulouse University Hospital, Toulouse, France and ^cFaculty of Health, University of Toulouse-Paul Sabatier, Toulouse, France

jonathan.curot@cnrs.fr
anais.servais@cnrs.fr
emmanuel.barbeau@cnrs.fr

doi:10.1017/S0140525X23000110, e362

Abstract

Déjà vu and involuntary autobiographical memories (IAM) can be induced by intracranial electric brain stimulation in epileptic patients, sometimes in the same individual. We suggest that there may be different types of IAM which should be taken into account and provide several ideas to test the hypothesis of a continuity between IAM and déjà vu phenomena.

Are déjà vu and involuntary autobiographical memories (IAM) based on the same continuum of retrieval processes? This is the core hypothesis proposed by the intellectually stimulating paper by Barzykowski and Moulin (B&M) which we will discuss specifically within the spectrum of déjà vu and IAM induced by intracranial electrical brain stimulation (iEBS) in epileptic patients.

There is an undeniable proximity between déjà vu and IAM with strong arguments presented in B&M's proposal. However, conceptual proximity does not mean mechanistic similarity. For instance, the authors emphasize that IAM are frequent, contrary to déjà vu, which are much rarer. One may wonder why phenomena sharing common mechanisms do not have a correlated prevalence. How this is accounted for is not yet clear. A better definition of IAM could clarify this issue. In fact, there are also inconsistencies in the reported frequency of IAM across studies. While some studies report high frequency, of up to 20 per day (Rasmussen & Berntsen, 2011), others report only one per week or per month (Moulin et al., 2014). Do such discrepancies mean that there are different types of IAM which are poorly differentiated in the literature at present?

B&M focus mainly on IAM that occur several times a day during mind wandering due to the associative nature of memory processes (Plimpton, Patel, & Kvilavilashvili, 2015). But how can IAM such as the memory of Tante Leonie's home in Combray which popped up in Marcel Proust's mind thanks to a madeleine be explained? Proust's IAM was sufficiently intrusive, unusual and disturbing for him to write 2,400 pages largely inspired by this phenomenon. And what about memories that are just as intense and intrusive with no associative trigger in the environment? We contend that such "pop-up" IAM are rare and do not correspond to "mind-wandering" IAM. These "pop-up" IAM seem to have a similar prevalence to that of the déjà vu experience reported in the healthy population. In fact, pop-up IAM not triggered by an environmental or goal-directed memory cue are the minority (approximately 15%) of the IAM experienced in the everyday life of healthy subjects (Ball & Little, 2006).

Distinguishing between "pop-up" and "mind-wandering" IAM (this terminology may need to be revised) may be all the more relevant as they do not seem to serve the same purpose. While reminiscing on autobiographical memory during mind-wandering episodes is assumed to play a fundamental role in memory consolidation (Wamsley, 2019) and possibly in personality construction, whether or not "pop-up" IAM play an adaptive or ecological role is unclear. Again, in terms of functionality, this role seems closer to déjà vu which B&M describe as the result of a metamemory error. "Pop-up" IAM may also lead to some level of surprise similar to déjà vu.

Déjà vu and experiential phenomena in epileptic patients, whether occurring spontaneously during seizures or after iEBS, offer informative snapshots into these highly subjective memory phenomena which can be phenomenologically very rich (Curot et al., 2017). They have an important role as complements in cognitive psychology experiments, in discussing the reality of theoretical models, and in obtaining insights into the physiological mechanisms underlying such phenomena.

Like the memory of Tante Leonie's home for Proust, epileptic subjects can also experience "pop-up" IAM after an EBS and get literally invaded by such memories, unexpectedly and involuntarily (for verbatim details, see Curot et al. [2017] and corresponding online database <https://figshare.com/s/923f93555a0ce51426e4> of >370 experiential phenomena reported in >110 patients; Bartolomei, Lagarde, Médina Villalon, McGonigal, & Benar,

2016). Their attention is suddenly projected toward their internal word, which could be interpreted as mental diplopia or a dissociation process in these patients. Patients often spontaneously describe a feeling of temporal and/or spatial projection into their own past. During these "pop-up" IAM, the intention to remember is absent, which may explain why it is such a surprise: There is no "madeleine" or environmental trigger. Interestingly, B&M mention that one of the limitations of the literature is that déjà vu and IAM are studied separately, while the EBS-induced déjà vu and IAM model allows studying these phenomena together in the same patients. Therefore, we propose using this approach to investigate a possible continuum between déjà vu and IAM. Based on the observations made in epileptic patients after iEBS, we suggest that there may be different types of IAM, including "pop-up" IAM, which may be closer to déjà vu than to mind-wandering IAM. We agree that déjà vu and IAM could be part of a continuum involving retrieval processes, but that this could be the case for a certain type of IAM only, those that pop up in the mind without any obvious memory cue.

To examine this hypothesis, we could: (1) Calculate the intra-individual correlation between the number of IAM and déjà vu induced by iEBS in the same brain region; (2) investigate the correlation between the prevalence of these "pop-up" IAM and experiential phenomena such as déjà vu which occur during spontaneous seizures, individually and in a group. If "pop-up" IAM and déjà vu experiences rely on shared underlying mechanisms, we expect a significant positive correlation between their prevalence, but no correlation with mind-wandering IAM; (3) iEBS are performed during intracranial EEG recordings, enabling trapping of the neural correlates of the fleeting moments that correspond to IAM and déjà vu (Barbeau et al., 2005; Bartolomei et al., 2012). Functional connectivity networks during each phenomenon could be compared to identify commonality and differences in connectivity patterns. Each network could also be compared to those of episodic memory and mind wandering to assess the specificity of "pop-up" IAM. (4) Finally, in 2017, we proposed a classification of the different types of memories induced by EBS according to their content and the current main concepts of long-term memory (Curot et al., 2017, see online database). This classification could be reviewed in light of the suggestions made by B&M to verify how they fit with the verbatim report of patients.

Financial support. None.


Competing interest. None.

References

- Ball, C. T., & Little, J. C. (2006). A comparison of involuntary autobiographical memory retrievals. *Applied Cognitive Psychology*, 20(9), 1167–1179. <https://doi.org/10.1002/acp.1264>
- Barbeau, E., Wendling, F., Régis, J., Duncan, R., Poncet, M., Chauvel, P., & Bartolomei, F. (2005). Recollection of vivid memories after perirhinal region stimulations: Synchronization in the theta range of spatially distributed brain areas. *Neuropsychologia*, 43(9), 1329–1337. <https://doi.org/10.1016/j.neuropsychologia.2004.11.025>
- Bartolomei, F., Barbeau, E. J., Nguyen, T., McGonigal, A., Régis, J., Chauvel, P., & Wendling, F. (2012). Rhinal-hippocampal interactions during déjà vu. *Clinical Neurophysiology*, 123(3), 489–495. <https://doi.org/10.1016/j.clinph.2011.08.012>
- Bartolomei, F., Lagarde, S., Médina Villalon, S., McGonigal, A., & Benar, C. G. (2016). The "Proust phenomenon": Odor-evoked autobiographical memories triggered by direct amygdala stimulation in human. *Cortex*, 90, 173–175. <https://doi.org/10.1016/j.cortex.2016.12.005>
- Curot, J., Busigny, T., Valton, L., Denuelle, M., Vignal, J. P., Maillard, L., ... Barbeau, E. J. (2017). Memory scrutinized through electrical brain stimulation: A review of 80 years of experiential phenomena. *Neuroscience and Biobehavioral Reviews*, 78, 161–177. <https://doi.org/10.1016/j.neubiorev.2017.04.018>

- Moulin, C. J. A., Souchay, C., Bradley, R., Buchanan, S., Karadoller, D. Z., & Akan, M. (2014). Déjà vu in older adults. In B. Schwartz & A. Brown (Eds.), *The tip of the tongue states and related phenomena* (pp. 281–304). Cambridge University Press.
- Plimpton, B., Patel, P., & Kvavilashvili, L. (2015). Role of triggers and dysphoria in mind-wandering about past, present and future: A laboratory study. *Consciousness and Cognition*, 33, 261–276. <https://doi.org/10.1016/j.concog.2015.01.014>
- Rasmussen, A. S., & Berntsen, D. (2011). The unpredictable past: Spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition*, 20(4), 1842–1846. <https://doi.org/10.1016/j.concog.2011.07.010>
- Wamsley, E. J. (2019). Memory consolidation during waking rest. *Trends in Cognitive Sciences*, 23(3), 171–173. <https://doi.org/10.1016/j.tics.2018.12.007>

Déjà vu and involuntary autobiographical memories as two distinct cases of familiarity in patients with Alzheimer’s disease

Joanna Gautier^a, Samuel Bulteau^{b,c},
Guillaume Chapelet^{d,e} and Mohamad El Haj^{a,d,f} 

^aLaboratoire de psychologie des Pays de la Loire, LPPL, Nantes Université, Univ Angers, Nantes, France; ^bClinical Investigation Unit 18, Department of Addictology and Psychiatry, CHU Nantes, Nantes, France; ^cMethodS in Patients-Centered Outcomes and Health Research (SPHERE), Institut National de la Santé et de la Recherche Médicale (INSERM), Nantes Université, Nantes, France; ^dClinical Gerontology Department, CHU Nantes, Nantes, France; ^eUniversité de Nantes, Inserm, TENS, The Enteric Nervous System in Gut and Brain Diseases, IMAD, Nantes, France and ^fInstitut Universitaire de France, Paris, France
Joanna.gautier@univ-nantes.fr
samuel.bulteau@chu-nantes.fr
guillaume.chapelet@chu-nantes.fr
mohamad.elhaj@univ-nantes.fr

doi:10.1017/S0140525X23000171, e363

Abstract

The continuum between involuntary autobiographical memories and déjà vu, as proposed by Barzykowski and Moulin, can be better defined by considering research on autobiographical retrieval in patients with Alzheimer’s disease. Although autobiographical retrieval in patients with Alzheimer’s disease can generally be associated with a sense of familiarity, involuntary retrieval can trigger an auto-noetic experience of retrieval in these patients.

Barzykowski and Moulin (B&M) provide an interesting theoretical framework for a large body of cognitive research studying phenomena that come to mind spontaneously and/or without any preceding intention to think about them. More specifically, B&M compare involuntary autobiographical memories and déjà vu regarding, among other factors, familiarity. Within this comparison, the authors propose that involuntary autobiographical memories may be considered as recollections of the personal past, whereas the “déjà vu” phenomenon rather triggers a stronger experience of familiarity.

We believe that the distinction between involuntary autobiographical memories and déjà vu, as proposed by B&M, can be enriched by considering studies in amnesia, especially in patients with Alzheimer’s disease. Research has demonstrated how, while

the typical retrieval of autobiographical memories in patients with Alzheimer’s disease is characterized by familiarity, involuntary autobiographical memories, at least as cued by sensory cues, can trigger a strong recollection experience in these patients. Within this view, familiarity can be considered as a metacognitive process, a feeling which is generated from the fluent processing of autobiographical information. This feeling, guiding patients with Alzheimer’s disease through their own phenomenological experience, can range from the general sense of familiarity (i.e., noetic experience of déjà vu) to the increased subjective experience of remembering (i.e., auto-noetic experience during involuntary retrieval).

Autobiographical retrieval in patients with Alzheimer’s disease is typically characterized by overgenerality, that is, a difficulty to retrieve unique memories situated in time and space (El Haj, Antoine, Nandrino, & Kapogiannis, 2015; El Haj, Boutoleau-Brettonnière, & Gallouj, 2020; El Haj, Moustafa, Gallouj, & Robin, 2019; El Haj, Roche, Gallouj, & Gandolphe, 2017b). This overgenerality can be observed regardless of methodology or memory distribution (Irish et al., 2011; Moses, Culpin, Lowe, & McWilliam, 2004). Autobiographical overgenerality in patients with Alzheimer’s disease can lead to a diminished subjective experience of retrieval. More specifically, overgenerality can lead to shift from an ability to mentally relive past events toward a general sense of familiarity that may be expressed by patients as a sense of “having experienced this before” (El Haj et al., 2015). This sense of familiarity can, somehow, mirror that triggered by “déjà vu.” In other words, patients with Alzheimer’s disease may tend to retrieve autobiographical memories under the lens of a “déjà vu” perspective, lacking the richness of contextual and phenomenological information, and, consequently, the auto-noetic experience of retrieval.

Although autobiographical retrieval in patients with Alzheimer’s disease can generally be associated with a sense of familiarity, involuntary retrieval can trigger an auto-noetic experience of retrieval in the patients. This assumption can be supported by the large body of research demonstrating how sensory cuing can trigger involuntary retrieval and, consequently, an enhanced recollective experience. For instance, research on music-evoked autobiographical memories has demonstrated how these memories can trigger, thanks to involuntary retrieval, an enhanced subjective experience of retrieval in patients with Alzheimer’s disease (El Haj, Fasotti, & Allain, 2012). The same thing can be said for odor-evoked autobiographical memories in patients with Alzheimer’s disease (El Haj, 2022; El Haj, Gandolphe, Gallouj, Kapogiannis, & Antoine, 2017a). Research has also demonstrated how visual cuing (i.e., exposure to nostalgic films) may promote involuntary autobiographical retrieval and a strong subjective experience in patients with Alzheimer’s disease (Rasmussen, Salgado, Daustrand, & Berntsen, 2021).

Taken together, while autobiographical retrieval in patients with Alzheimer’s disease can be associated with a general sense of familiarity, involuntary retrieval can trigger an auto-noetic subjective experience during retrieval. Thus, the continuum between involuntary autobiographical memories and déjà vu, as proposed by B&M, can be better defined by considering research on autobiographical retrieval in patients with Alzheimer’s disease. Familiarity may be considered as a cornerstone to the foundation of the distinction between déjà vu and involuntary retrieval, not only in the general population, as proposed by B&M, but also in amnesia.

Financial support. None.

Competing interest. None.

References

- El Haj, M. (2022). Odor-evoked autobiographical memory in Alzheimer's disease? *Archives of Clinical Neuropsychology*, 37(2), 513–520. <https://doi.org/10.1093/arclin/acab074>
- El Haj, M., Antoine, P., Nandrino, J. L., & Kapogiannis, D. (2015). Autobiographical memory decline in Alzheimer's disease, a theoretical and clinical overview. *Ageing Research Reviews*, 23(Pt B), 183–192. <https://doi.org/10.1016/j.arr.2015.07.001>
- El Haj, M., Boutoleau-Bretonnière, C., & Gallouj, K. (2020). The past as seen by women and men with Alzheimer disease: Sex differences in autobiographical memory. *Alzheimer Disease & Associated Disorders*, 34(2), 170–174. <https://doi.org/10.1097/WAD.0000000000000363>
- El Haj, M., Fasotti, L., & Allain, P. (2012). The involuntary nature of music-evoked autobiographical memories in Alzheimer's disease. *Consciousness and Cognition*, 21(1), 238–246. <https://doi.org/10.1016/j.concog.2011.12.005>
- El Haj, M., Gandalphe, M. C., Gallouj, K., Kapogiannis, D., & Antoine, P. (2017a). From nose to memory: The involuntary nature of odor-evoked autobiographical memories in Alzheimer's disease. *Chemical Senses*, 43(1), 27–34. <https://doi.org/10.1093/chemse/bjx064>
- El Haj, M., Moustafa, A. A., Gallouj, K., & Robin, F. (2019). Visual imagery: The past and future as seen by patients with Alzheimer's disease. *Consciousness and Cognition*, 68, 12–22. <https://doi.org/https://doi.org/10.1016/j.concog.2018.12.003>
- El Haj, M., Roche, J., Gallouj, K., & Gandalphe, M. C. (2017b). Autobiographical memory compromise in Alzheimer's disease: A cognitive and clinical overview. *Geriatric Et Psychologie Neuropsychiatrie Du Vieillessement*, 15(4), 443–451. <https://doi.org/10.1684/pnv.2017.0704> (Difficultés de rappel autobiographique dans la maladie d'Alzheimer: les enjeux cognitifs et cliniques.)
- Irish, M., Hornberger, M., Lah, S., Miller, L., Pengas, G., Nestor, P. J., ... Piguet, O. (2011). Profiles of recent autobiographical memory retrieval in semantic dementia, behavioural-variant frontotemporal dementia, and Alzheimer's disease. *Neuropsychologia*, 49(9), 2694–2702. <https://doi.org/10.1016/j.neuropsychologia.2011.05.017>
- Moses, A., Culpin, V., Lowe, C., & McWilliam, C. (2004). Overgenerality of autobiographical memory in Alzheimer's disease. *The British Journal of Clinical Psychology*, 43(Pt 4), 377–386. <https://doi.org/10.1348/0144665042388964>
- Rasmussen, K. W., Salgado, S., Daustrand, M., & Berntsen, D. (2021). Using nostalgia films to stimulate spontaneous autobiographical remembering in Alzheimer's disease. *Journal of Applied Research in Memory and Cognition*, 10(3), 400–411. <https://doi.org/https://doi.org/10.1016/j.jarmac.2020.11.001>

Involuntary memories are not déjà vu

Sami Gülgöz  and Irem Ergen

Psychology Department, Koç University, Rumeli Feneri Yolu, Sariyer, Istanbul, Turkey

sgulgöz@ku.edu.tr

iergen17@ku.edu.tr

kuram.ku.edu.tr

doi:10.1017/S0140525X2300002X, e364

Abstract

The proposed framework can benefit from integrating predictive processing into the explanation of déjà vu which corresponds to interrupted prediction. Déjà vu is also accompanied by familiarity. However, considerable ambiguity is inherent in familiarity, which necessitates elaboration of this construct. Research findings on involuntary autobiographical memories and déjà vu show discrepancies, and clustering these constructs can be counterproductive for research.

We believe that the authors have proposed a comprehensive approach to integrating the process of memory retrieval to account for both involuntary autobiographical memories (IAMs) and déjà vu experiences. In the current commentary,

we focus on four points that may contribute to the discussion. These are: (a) The need to integrate predictive processing into the explanatory framework for déjà vu, (b) the need to further specify the concepts of weak activation and familiarity, (c) the findings conflicting with the idea of common origins for déjà vu and IAM, and (d) the possible detriments of clustering constructs.

In recent years, the concept of prediction emerged with high impact to the extent that some researchers consider the study of prediction a new paradigm (Hutchinson & Barrett, 2019). Predictive processing emerges in areas of study from perception and imagery (Grush, 2004) to visual recognition (Bar et al., 2006) and decision-making (Doya, 2008). Cleary and Claxton (2018) suggested that déjà vu, too, could be a function of predictive processing. Even in novel circumstances, top-down processes seek contextual cues to past contexts to activate relevant past experiences. We use these past experiences or schema representations to predict current experiences. The déjà vu experience may happen when this prediction is erroneous. The mismatch between the prediction and the actuality would dissipate the fleeting feeling of familiarity, and the individual reports a déjà vu. Cleary et al. (2021) emphasized the familiarity component and showed that not all déjà vu experiences are accompanied by a feeling of prediction, but the feeling of familiarity typically characterizes them. Predictive processing studies rely on self-report making them difficult to observe because prediction remains implicit in daily life, whereas déjà vu is defined by awareness.

Researchers employing event-related potentials to study familiarity more directly obtained varying results. For example, familiar faces show N250, whereas participants' own faces do not (Sommer et al., 2021; but also Wiese et al., 2022). Caharel and Rossion (2021) showed that when long-term familiarity exists, the response is much sooner. Leynes and Upadhyay (2022) used familiar words and showed that responses changed depending on the context. Therefore, one must use caution when considering familiarity as a singular construct.

The concepts of familiarity and weak activation form the main tenets of the authors' hypotheses. These constructs require further elaboration because several factors may be influential, for example, insufficient cues, the inadequacy of cues, threshold for activation, or activation level. In addition, weak familiarity may arise because memory cells or ensembles have weak connectivity, cell clusters are partially damaged, or the representation is temporarily inaccessible (Josselyn & Tonegawa, 2020). If familiarity is a function of contextual cues, there must be a sufficient level of cues to trigger familiarity, but the configuration of these cues may not match the activated representation. We have also seen that schema activation and remembering specific memories cannot be distinguished easily in recalling past events (Ece & Gülgöz, 2021). In déjà vu, it is likely that the cues activate a schema rather than a specific event, and thus, an actual memory may not be available. In IAMs, on the other hand, what is remembered is a previously experienced specific event. It is also possible that déjà vu results from the hippocampus being unable to reinstate the original memory representation or to reconstruct it (Barry & Maguire, 2019; Frankland, Josselyn, & Köhler, 2019). With IAMs, on the other hand, memory is reinstated without any deliberation.

Another point is the discrepancies between IAM and déjà vu findings. Berntsen (2010) proposed that IAM and voluntary recall share the same encoding and storage principles with four claims. First, the IAM recall experience is universal, and they are observed

in most individuals (Ball & Little, 2006; Rubin & Berntsen, 2009), whereas Brown (2003) reported the rate of people experiencing déjà vu at least once in their lifetime to be 67%. Another discrepancy is the decrease in the frequency of déjà vu experience with age. No such decrease is observed in IAM experience (Berntsen et al., 2015). One would expect that the likelihood of having a déjà vu would increase with experience and the possibility of an associative match, according to configural similarity and familiarity theories. Berntsen (2010) also proposed that IAM recall is frequent, with around 20 memories per day on average (Rasmussen & Berntsen, 2011), but déjà vu experience is very rare, a few times a year (Brown, 2003). If familiarity increases déjà vu experience, then déjà vu experience should be more frequent because our environment is full of cues overlapping with our experience (Cleary et al., 2012; Cleary & Claxton, 2018). Berntsen (2010) argued that involuntary memories belong to the same memory system as voluntary ones. However, IAMs differ in retrieval processes (Barzykowski & Staugaard, 2016). Déjà vu phenomenon is difficult to categorize within the memory system because it does not include any key processes observed in memory, such as encoding, maintenance, or retrieval. In déjà vu, what is lacking is not “access to content” (see Fig. 1 in the main article), but the content itself. Conversely, there is an actual recall of “content” in IAM. Therefore, déjà vu is merely a spontaneous and false belief of familiarity.

Although trying to explicate memory mechanisms within a single framework and to have a general model of retrieval is valuable, clustering qualitatively distinct processes as on a continuum may blind researchers to nuances in the underlying processes. There seem to be similar problems with other constructs. For example, false memory is an overarching term that subsumes memory implantation, recovered memories, imagination inflation, misinformation effect, and the Deese, Roediger and McDermott (DRM) task although there is evidence that they may be products of different mechanisms. Patihis, Frenda, and Loftus (2018) studied the correlations between different false memory paradigms and observed no relationship. In order to avoid such pitfalls, we should focus on particular mechanisms rather than constructing general frameworks, at least for now.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

- Ball, C. T., & Little, J. C. (2006). A comparison of involuntary autobiographical memory retrievals. *Applied Cognitive Psychology, 20*(9), 1167–1179.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., ... Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences, 103*(2), 449–454.
- Barry, D. N., & Maguire, E. A. (2019). Remote memory and the hippocampus: A constructive critique. *Trends in Cognitive Sciences, 23*(2), 128–142.
- Barzykowski, K., & Staugaard, S. R. (2016). Does retrieval intentionality really matter? Similarities and differences between involuntary memories and directly and generatively retrieved voluntary memories. *British Journal of Psychology, 107*(3), 519–536.
- Berntsen, D. (2010). The unbidden past: Involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science, 19*(3), 138–142.
- Berntsen, D., Rubin, D. C., & Salgado, S. (2015). The frequency of involuntary autobiographical memories and future thoughts in relation to daydreaming, emotional distress, and age. *Consciousness and Cognition, 36*, 352–372.
- Brown, A. S. (2003). A review of the déjà vu experience. *Psychological Bulletin, 129*, 394–413.
- Caharel, S., & Rossion, B. (2021). The N170 is sensitive to long-term (personal) familiarity of a face identity. *Neuroscience, 458*, 244–255.

- Cleary, A. M., Brown, A. S., Sawyer, B. D., Nomi, J. S., Ajoku, A. C., & Ryals, A. J. (2012). Familiarity from the configuration of objects in 3-dimensional space and its relation to déjà vu: A virtual reality investigation. *Consciousness and Cognition, 21*(2), 969–975.
- Cleary, A. M., & Claxton, A. B. (2018). Déjà vu: An illusion of prediction. *Psychological Science, 29*(4), 635–644.
- Doya, K. (2008). Modulators of decision making. *Nature Neuroscience, 11*(4), 410–416.
- Ece, B., & Gülgöz, S. (2021). Autobiographical memory for repeated events: Remembering our vacations. *Journal of Cognitive Psychology, 33*(1), 1–11.
- Frankland, P. W., Josselyn, S. A., & Köhler, S. (2019). The neurobiological foundation of memory retrieval. *Nature Neuroscience, 22*(10), 1576–1585.
- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences, 27*(3), 377–396.
- Hutchinson, J. B., & Barrett, L. F. (2019). The power of predictions: An emerging paradigm for psychological research. *Current Directions in Psychological Science, 28*(3), 280–291.
- Josselyn, S. A., & Tonegawa, S. (2020). Memory engrams: Recalling the past and imagining the future. *Science, 367*(6473), eaaw4325.
- Leynes, P. A., & Upadhyay, T. (2022). Context dissociations of the FN400 and N400 are evidence for recognition based on relative or absolute familiarity. *Brain and Cognition, 162*, 105903.
- Patihis, L., Frenda, S. J., & Loftus, E. F. (2018). False memory tasks do not reliably predict other false memories. *Psychology of Consciousness: Theory, Research, and Practice, 5*(2), 140.
- Rasmussen, A. S., & Berntsen, D. (2011). The unpredictable past: Spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition, 20*(4), 1842–1846.
- Rubin, D. C., & Berntsen, D. (2009). The frequency of voluntary and involuntary autobiographical memories across the life span. *Memory & Cognition, 37*(5), 679–688.
- Sommer, W., Stapor, K., Kończak, G., Kotowski, K., Fabian, P., Ochab, J., ... Ślusarczyk, G. (2021). The N250 event-related potential as an index of face familiarity: A replication study. *Royal Society Open Science, 8*(6), 202356.
- Wiese, H., Hobden, G., Siilbek, E., Martignac, V., Flack, T. R., Ritchie, K. L., ... Burton, A. M. (2022). Familiarity is familiarity is familiarity: Event-related brain potentials reveal qualitatively similar representations of personally familiar and famous faces. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(8), 1144–1164.

Deconstructing spontaneous expressions of memory in dementia

Muireann Irish^{a,b} 

^aSchool of Psychology, The University of Sydney, Camperdown, NSW, Australia and ^bBrain and Mind Centre, The University of Sydney, Camperdown, NSW, Australia

muireann.irish@sydney.edu.au

<https://www.forefrontresearch.org/professor-muireann-irish/>

doi:10.1017/S0140525X23000018, e365

Abstract

Dementia syndromes offer a unique opportunity to clarify some of the component processes of spontaneous expressions of memory proposed by the Barzykowski and Moulin model. By considering the model through the lens of memory disorders, I outline several important extensions to progress our understanding of these spontaneous cognitive phenomena.

A nascent topic in cognitive neuroscience is the relatively seamless manner by which humans transition between externally evoked and internally driven expressions of memory. Barzykowski and Moulin (B&M) present a new framework under which seemingly divergent manifestations of spontaneous cognition might be accommodated within the same conceptual space. Two putative

expressions of spontaneous memory – involuntary autobiographical memory and déjà vu – are conceptualised as interlinked processes that vary in terms of their content and associated phenomenology. By situating these processes along a representational continuum, the framework provides an interesting testbed for future empirical studies to capture variable manifestations of spontaneous cognition in health and disease.

Despite its broad appeal, several theoretical aspects warrant further development, particularly in relation to disorders of memory. First, the model proposes that the threshold for familiarity is intimately linked to the retrieval of contextual specifics. Accordingly, deterioration of episodic or semantic fluency signals should impede re-instantiation of the corresponding neural activation of the original experience, resulting in a subjective feeling of familiarity (reviewed by Renoult, Irish, Moscovitch, & Rugg, 2019). Degradation of the contextual memory trace and accompanying recollective experience is a transdiagnostic feature of dementia (Irish, 2023; Irish, Piquet, & Hodges, 2012). On formal tests of autobiographical memory, participants with Alzheimer's disease display a content-specific vulnerability, biasing the subjective experience towards a feeling of familiarity in the absence of detailed recollection (reviewed by Irish, 2023). Thus, for deliberate forms of autobiographical retrieval, at least, the model's predictions of a coupling between content and phenomenology appear well supported.

Importantly, however, the B&M model focuses on *involuntary* expressions of memory that arise relatively unbidden or without deliberate intention. For involuntary autobiographical memories (IAM), the central premise is that a content-addressable memory springs to mind fully formed, reflecting an automatic bottom-up retrieval process in response to environmental cues. Few studies have explored the frequency, content or phenomenology of IAMs in dementia; however, studies of self-referential forms of spontaneous cognition provide initial insights (O'Callaghan & Irish, 2018). Participants with Alzheimer's disease have been shown to generate significantly fewer incidences of mind wandering relative to healthy older adults during ongoing performance of a sustained attention task (Gyrkovics, Balota, & Jackson, 2018). In contrast, work from our lab indicates that under conditions of low cognitive demand, some types of self-referential mind wandering can be achieved in Alzheimer's disease (O'Callaghan, Shine, Hodges, Andrews-Hanna, & Irish, 2019). These conflicting findings suggest that the way in which we probe spontaneous cognition significantly influences the resultant content and subjective experience.

How, then, might incidental cues in the external environment give rise to somewhat aberrant expressions of memory, such as déjà vu? Here, Barzykowski and Moulin's (B&M) model diverges from previous literature by intimating that a loss of contextual detail could *potentiate* a feeling of déjà vu. Contexts that are repeatedly encountered might trigger a schema-driven sense of familiarity, which lacks the spatiotemporal specificity required for fully fledged episodic recollection (Irish & Vatansever, 2020). In contexts where an appropriate detail threshold is not met and feelings of familiarity are not corroborated with sufficient detail, the ensuing representation might shift from a content-addressable memory to the subjective experience of déjà vu. Based on the B&M model, one might be tempted to predict elevated levels of déjà vu in dementia populations, reflecting a mismatch between preserved familiarity for an incidental cue or context in the absence of the comparator content from episodic memory. Empirical research on déjà vu in dementia, however,

is extremely scant, given the inherent difficulties in operationalising and measuring a rare occurrence that requires high levels of metacognitive insight for its detection.

Finally, we must consider the endogenous processes by which spontaneous cognition emerges in response to incidental external triggers. Recent theoretical frameworks place incidental external triggers at the cornerstone of involuntary quirks of memory and their natural occurrence in everyday life. The spontaneous retrieval deficit hypothesis argues that spontaneous manifestations of autobiographical memory are disproportionately compromised in dementia due to an inability to endogenously generate appropriate content in response to stimuli encountered in the environment (Kvavilashvili, Niedzwienska, Gilbert, & Markostamou, 2020). This hypothesis emphasises the importance of automatic bottom-up retrieval of cue-driven representations as an early marker of Alzheimer's disease, but one which seems minimally affected in healthy aging (see also Irish, Goldberg, Alaeddin, O'Callaghan, & Andrews-Hanna, 2019). A marked attenuation of spontaneous phenomena, including IAMs and déjà vu, would therefore be predicted in any population where the endogenous generation of mental content is impaired (O'Callaghan & Irish, 2018).

While B&M's model offers an interesting glimpse into potential mechanisms driving spontaneous quirks of memory, its application in neurocognitive disorders raises several challenges and opportunities for future research. Further iterations of the model might explore the nature of external cues in determining the frequency and readout of spontaneous memory. Efforts to delineate how different levels of cue specificity influence the perceptual and conceptual fluency signals that drive familiarity processes in dementia will also be essential. Ultimately, by incorporating synergies and discontinuities from clinical populations, the model can progress towards a more integrative account.


Financial support. Muireann Irish is supported by an Australian Research Council Discovery Project (DP220100663).

Competing interest. None.

References

- Gyrkovics, M., Balota, D. A., & Jackson, J. D. (2018). Mind-wandering in healthy aging and early stage Alzheimer's disease. *Neuropsychology*, 32(1), 89–101.
- Irish, M. (2023). Autobiographical memory in dementia syndromes – an integrative review. *Wiley Interdisciplinary Reviews. Cognitive Science*, 14, e1630. doi: 10.1002/wcs.1630
- Irish, M., Goldberg, Z. L., Alaeddin, S., O'Callaghan, C., & Andrews-Hanna, J. R. (2019). Age-related changes in the temporal focus and self-referential content of spontaneous cognition during periods of low cognitive demand. *Psychological Research*, 83(4), 747–760. doi: 10.1007/s00426-018-1102-8
- Irish, M., Piquet, O., & Hodges, J. R. (2012). Self-projection and the default network in frontotemporal dementia. *Nature Reviews. Neurology*, 8(3), 152–161. doi: 10.1038/nrneuro.2012.11 [pii]. 10.1038/nrneuro.2012.11 [doi]
- Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, 32, 43–49.
- Kvavilashvili, L., Niedzwienska, A., Gilbert, S. J., & Markostamou, I. (2020). Deficits in spontaneous cognition as an early marker of Alzheimer's disease. *Trends in Cognitive Sciences*, 24(4), 285–301. doi: 10.1016/j.tics.2020.01.005
- O'Callaghan, C., & Irish, M. (2018). Candidate mechanisms of spontaneous cognition as revealed by dementia syndromes. In K. Christoff & K. C. Fox (Eds.), *The Oxford handbook of spontaneous thought: Mind-wandering, creativity, and dreaming* (pp. 493–508). Oxford University Press.
- O'Callaghan, C., Shine, J. M., Hodges, J. R., Andrews-Hanna, J. R., & Irish, M. (2019). Hippocampal atrophy and intrinsic brain network dysfunction relate to alterations in mind wandering in neurodegeneration. *Proceedings of the National Academy of Sciences of the USA*, 116(8), 3316–3321. doi: 10.1073/pnas.1818523116
- Renoult, L., Irish, M., Moscovitch, M., & Rugg, M. D. (2019). From knowing to remembering: The semantic-episodic distinction. *Trends in Cognitive Sciences*, 23(12), 1041–1057.

Distinguishing involuntary autobiographical memories and déjà vu experiences: Different types of cues and memory representations?

Lia Kvavilashvili  and Ioanna Markostamou

Psychology Division, School of Life and Medical Sciences, University of Hertfordshire, Hatfield, Hertfordshire, UK

l.kvavilashvili@herts.ac.uk

i.markostamou@herts.ac.uk

<http://www.liakvavilashvili.com/>

[https://researchprofiles.herts.ac.uk/portal/en/persons/lia-kvavilashvili\(48d7e4e1-c235-4f92-9f13-8832aae8f2f2\).html](https://researchprofiles.herts.ac.uk/portal/en/persons/lia-kvavilashvili(48d7e4e1-c235-4f92-9f13-8832aae8f2f2).html)

[https://researchprofiles.herts.ac.uk/portal/en/persons/ioanna-markostamou\(4747a504-fa5e-4035-9652-bb4abdb6b38c\).html](https://researchprofiles.herts.ac.uk/portal/en/persons/ioanna-markostamou(4747a504-fa5e-4035-9652-bb4abdb6b38c).html)

doi:10.1017/S0140525X23000080, e366

Abstract

Although involuntary autobiographical memories (IAMs) and déjà vu have important shared characteristics, in this commentary, we focus on potential differences that may question the argument that two phenomena lie on a continuum. We propose that differences in their frequency and auto-noetic consciousness could be explained by different types of cues and memory representations involved in experiencing IAMs and déjà vu.

The target article proposes that the basic retrieval mechanisms involved in involuntary autobiographical memories (IAMs) and déjà vu may be more similar than previously thought. While we do not disagree with this premise and find the proposed framework useful in stimulating future research, in this commentary, we focus on two key aspects that differentiate IAMs from déjà vu – the temporal orientation of the phenomenological experience of the two phenomena (while IAMs involve recalling and re-living past experiences, déjà vu is characterised by a feeling of re-living the present moment) and the frequency of their occurrence (while IAMs occur very frequently, déjà vu occurs very rarely). We argue that these differences may be explained by two crucial differences in the memory processes involved: The triggering conditions and the types of events being recalled.

In the target article, one of the key similarities between IAMs and déjà vu refers to auto-noetic consciousness – the sense of re-living a *past* experience or traveling back in time when recalling a past event. While this definitely applies to IAMs, in cases of déjà vu, a person experiences an intense feeling of familiarity that they have been in the current situation before, in other words, they experience a sense of re-living the *present* moment as if it had already occurred (without recalling a past event). This means that while IAMs refer to a form of conscious *recollective memory*, déjà vu experiences refer to an unusual form of *recognition memory* where the intense feeling of familiarity never gets resolved (unlike a “butcher-on-the-bus” example). This core phenomenological difference raises some doubts about the two phenomena being on a continuum. Moreover, if the two phenomena rely on the same basic recall mechanisms, then why do IAMs occur multiple times a day (Gardner & Ascoli, 2015) while déjà vu experiences are so rare? Below we address these issues by examining the

role of triggers in eliciting the two phenomena and the types of memory representations that may be activated.

It is well established that simple and easily identifiable cues, which map onto a previously experienced – and recallable – event, trigger IAMs (Mace, 2004; Schlagman, Kvavilashvili, & Schulz, 2007). Cues can be objects (a broken window), individuals (an old lady) or places (a street corner), or they can have a verbal form (a written word or a conversation) (Berntsen & Hall, 2004). Such simple cues have been successfully used to elicit IAMs in laboratory vigilance tasks (Schlagman & Kvavilashvili, 2008; Vannucci et al., 2015). The literature reviewed in the target article, however, suggests that déjà vu experiences are elicited by being in a particular situation or scene that features an ensemble of multiple (and possibly subtle) cues, that create a sense of re-living this very same situation in the present moment as something that had already occurred in the past. Thus, while an easily identifiable and rather isolated cue in the environment can trigger an IAM and transport us back to some past event that occurred in a very different context (e.g., seeing a candle on TV eliciting a memory of romantic dinner last weekend), in déjà vu, the feeling of familiarity is triggered by a current situation as a whole, with multiple contextual cues present. These differences in cueing conditions make it highly unlikely that standard vigilance tasks using incidental verbal cues to study IAMs will elicit instances of déjà vu as suggested in the target article. These differences, however, are not sufficient to explain why déjà vu experiences are so rare compared to the frequent occurrences of IAMs.

The second and perhaps more critical factor as suggested in the target paper is that a particular constellation of cues in a scene or situation only partially activates traces of some previously encountered scene or situation. This may occur if there is only a subtle resemblance between the current and the previous scene, as demonstrated by Cleary and colleagues’ work described in the target paper. Consider, for example, that you walk down a street in a new city and see a beautiful glass ball in an art gallery window. The colour and the texture of this ball may elicit a fond memory of swimming with dolphins in Hawaii. However, a strong déjà vu could be experienced if the situation of encountering this glass ball has some resemblance to a situation of seeing a beach ball in a sports shop window when walking along a street in London some time ago. The resemblance between the balls is too weak to activate a memory of seeing the beach ball in London; however, there may be several other subtle overlapping cues such as the presence of an ice-cream stall near both shops or a strong wind blowing in your face that may create, in this particular constellation, this intense feeling of familiarity and re-living the present moment.

Importantly, this hypothetical example suggests that memory traces that get partially activated in déjà vu are not fragments of sensory-perceptual experiences that have become part of one’s autobiographical knowledge system (Conway, 2009), and can be easily recalled in response to simple cues. Instead, partially activated representations of previously encountered scenes may be part of one’s long-term perceptual-representation system. Although this system was originally proposed to explain implicit memory phenomena (Tulving & Schacter, 1990), there is growing evidence from multiple strands of research (Brewin, 2014) that people can retain a large amount of incidentally encountered information (scenes, stimuli), that may even get accessed consciously under certain circumstances. This has been demonstrated by research using SenseCam pictures to cue memories (Sellen et al., 2007) and in studies testing recognition memory for

hundreds and even thousands of pictures and scenes (Delorme, Poncet, & Fabre-Thorpe, 2018; Standing, 1973). It is therefore possible that IAMs and déjà vu experiences differ also in types of representations that get activated. One interesting prediction that emerges from this proposition is that participants scoring high in such experiments may be prone to more frequent self-reported déjà vu.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interest. None.

References

- Berntsen, D., & Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Memory and Cognition*, 32, 789–803.
- Brewin, C. R. (2014). Episodic memory, perceptual memory, and their interaction: Foundations for a theory of posttraumatic stress disorder. *Psychological Bulletin*, 140, 69–97. <http://dx.doi.org/10.1037/a0033722>
- Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, 47, 2305–2313. <https://doi.org/10.1016/j.neuropsychologia.2009.02.003>
- Delorme, A., Poncet, M., & Fabre-Thorpe, M. (2018). Briefly flashed scenes can be stored in long-term memory. *Frontiers in Neuroscience*, 12, 688. doi:10.3389/fnins.2018.00688
- Gardner, R. S., & Ascoli, G. A. (2015). The natural frequency of human prospective memory increases with age. *Psychology and Aging*, 30, 209–219. <https://doi.org/10.1037/a0038876>
- Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract cuing: The Proustian view is incorrect. *Applied Cognitive Psychology*, 18, 893–899. <https://doi.org/10.1002/acp.1020>
- Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory and Cognition*, 36, 920–932. doi:10.3758/MC.36.5.920
- Schlagman, S., Kvavilashvili, L., & Schulz, J. (2007). Effects of age on involuntary autobiographical memories. In J. H. Mace (Ed.), *Involuntary memory* (pp. 87–112). Blackwell.
- Sellen, A., Fogg, A., Aitken, M., Hodges, S., Rother, C., & Wood, K. (2007). Do logging technologies support memory for the past? An experimental study using SenseCam. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 81–90). Association for Computing Machinery. doi:10.1145/1240624.1240636
- Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, 25, 207–222. doi:10.1080/14640747308400340
- Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, 247, 301–306. doi:10.1126/science.2296719
- Vannucci, M., Pelagatti, C., Hanczakowski, M., Mazzoni, G., & Paccani, C. R. (2015). Why are we not flooded by involuntary autobiographical memories? Few cues are more effective than many. *Psychological Research*, 79, 1077–1085. <https://doi.org/10.1007/s00426-014-0632-y>

A rational analysis and computational modeling perspective on IAM and déjà vu

Justin Li^a, Steven Jones^b and John Laird^b 

^aDepartments of Cognitive Science and Computer Science, Occidental College, Los Angeles, CA, USA and ^bCenter for Integrated Cognition, Ann Arbor, MI, USA
justinnhli@oxy.edu
steven.jones@cic.iqmri.org
john.laird@cic.iqmri.org
<https://www.oxy.edu/academics/faculty/justin-li>
<https://scijones.github.io/>
<https://laird.engin.umich.edu/>

doi:10.1017/S0140525X23000092, e367

Abstract

The proposed memory architecture by Barzykowski and Moulin is compelling, and could be improved by incorporating a rational analysis of the functional roles of involuntary autobiographical memory and déjà vu. Additionally, modeling these phenomena computationally would remove ambiguities from the proposal. We provide examples of past work that illustrate how the phenomena may be described more precisely.

The target article by Barzykowski and Moulin (B&M) argues that involuntary autobiographical memory (IAM) and déjà vu are the result of an integrated system of memory and that they naturally arise from recognition and memory retrieval processes. While we agree with this stance, we find the proposed memory architecture lacking in two respects. First, the focus of the authors on the phenomenology of IAM and déjà vu neglects the functional role of these phenomena, the justification for why they might exist, and how they might be used by agents. Second, the memory architecture, as summarized in Figure 1, is only imprecisely specified, leaving room for alternate theories, potential inconsistencies, and omitted details. Here, we consider how performing a rational analysis of, and building computational models of, IAM and déjà vu can mitigate these problems. We use our work on how IAM can support prospective memory as an example of addressing both problems (Li & Laird, 2015), then extend that reasoning to the familiarity judgments that underlie déjà vu.

The rational analysis framework assumes that cognitive processes are optimally adapted to the functional goals of the agent, while subject to ecological constraints and limits on biological and cognitive resources (Anderson, 1990; Lieder & Griffiths, 2020). For memory, we take its primary function to be to “bring past experience to bear on present action” (Anderson, 1994), operating within a small working memory capacity, fixed bandwidth to long-term memory, and other cognitive constraints. This serves as the starting point for understanding the functional role of phenomena such as IAM.

A hypothesis about the structure of memory can be tested via its implementation in a computational model. These models force researchers to be precise in their definitions of the computational representations and processes that underlie their theories and ensure that hypothesized theoretical models of memory are consistent both internally and with broader theories of cognition. This is particularly true in cognitive architectures such as ACT-R (Anderson, 2007), which integrate multiple cognitive processes and potential neural correlates into a single system. This enables the evaluation of their combined performance across multiple tasks, thus ensuring that a hypothesis is compatible with the same mechanisms used to model other phenomena.

Consider, for example, the hypothesis that IAM is the result of automatic matching of sensory and abstract cues with items in memory in order to “quickly rais[e] pertinent information to consciousness without effort” (target article, sect. 4, para. 3). A rational analysis of IAM would start by considering the limits of deliberate retrieval and situations where those limits are exceeded. One such situation is in the prospective memory for future goals, when there may not be the intention to initiate a deliberate retrieval. For example, if one were previously asked to pass a message to a colleague, nothing about seeing the colleague later in the day would necessarily prompt a deliberate retrieval to bring that task to mind, especially if the encounter is otherwise routine. It

is in this context that IAM provides a functional benefit, and indeed, this is known as spontaneous retrieval in the prospective memory literature and is one of several possible strategies for achieving such a goal (McDaniel & Einstein, 2007). This application of rational analysis showcases how IAM can play a role in problem solving: Beyond the passive role suggested in the target article, people can learn to take advantage of IAM to reduce cognitive load.

We have implemented a spontaneous retrieval mechanism in a cognitive architecture, modeled its use in prospective memory, and shown that the conditions under which it succeeds qualitatively resemble results from human experiments (Li & Laird, 2015). More than that, the model requires a fully specified theory of how IAM arises and how it interacts with other memory processes such as deliberate retrieval; in our case, to prioritize problem solving, involuntary retrieval only occurs when no deliberate retrievals are taking place. This decision, which follows from the assumption that memory is used to support goal-driven behavior, suggests an explanation for why researchers have found that IAMs occur most commonly during “relaxed or non-focused state[s] of awareness” and how “being focused would inhibit the activation of knowledge units that are inconsistent with the individual’s current goals” (Berntsen, 2008; quoting Mandler, 1994). In this case, the computational model built on the rational analysis framework aligns with the psychology literature.

As for déjà vu, although we know of no existing theory of its functional role in cognition, we agree that it results from false positives in familiarity judgments. Familiarity judgments – or at least its simplest form, recognition – have long been a subject of study via rational analysis, with recognition probability following the optimal Bayesian solution (Shiffrin & Steyvers, 1997). That familiarity is faster than recall allows it to be used to guide the strategic search for knowledge, as per the cognitive-heuristic account of metamemory (Schwartz & Metcalfe, 2011). We have implemented recognition judgments in a cognitive architecture and used it to trigger deliberate retrievals, which led to situations where a false-positive recognition resulted in retrieval failure (Li, Derbinsky, & Laird, 2012). Although we have not modeled déjà vu explicitly, the retrieval failure could suggest that recognition was “implausible.” This interaction between deliberate retrieval and plausibility was not explored in the target article, nor how familiarity and recollection interact with each other over time.

In sum, the target article by B&M presents a compelling proposal for how IAM and déjà vu arise. However, their description is missing details that would clarify the relationships between memory mechanisms and could be improved by accounting for the functionality of these phenomena. Applying the rational analysis framework, and considering how the proposed system may be modeled computationally, would resolve these issues.

Financial support. The authors did not receive any outside funding for this work.


Competing interest. None.

References

- Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.
- Anderson, N. H. (1994). Functional memory. In *A functional theory of cognition* (pp. 363–394). Psychology Press.
- Berntsen, D. (2008). Involuntary autobiographical memories: Speculations, findings, and an attempt to integrate them. In J. H. Mace (Ed.), *Involuntary memory* (pp. 20–49). Blackwell Publishing. <https://doi.org/10.1002/9780470774069.ch2>

- Li, J., Derbinsky, N., & Laird, J. E. (2012). Functional interactions between memory and recognition judgments. In *Proceedings of the 26th AAAI conference on artificial intelligence* (pp. 228–234). AAAI Press.
- Li, J., & Laird, J. E. (2015). Spontaneous retrieval for prospective memory: Effects of encoding specificity and retention interval. In *Proceedings of the 13th international conference on cognitive modeling* (pp. 142–147). University of Gronigen Press.
- Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43, e1. <https://doi.org/10.1017/S0140525X1900061X>
- Mandler, G. (1994). Hypermnesia, incubation, and mind popping: On remembering without really trying. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance 15: Conscious and nonconscious information processing* (pp. 3–33). The MIT Press.
- McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Sage.
- Schwartz, B. L., & Metcalfe, J. (2011). Tip-of-the-tongue (TOT) states: Retrieval, behavior, and experience. *Memory & Cognition*, 39(5), 737–749. <https://doi.org/10.3758/s13421-010-0066-8>
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM – retrieving effectively from memory. *Psychonomic Bulletin and Review*, 4(2), 145–166.

Are involuntary autobiographical memory and déjà vu cognitive failures?

John H. Mace 

Department of Psychology, Eastern Illinois University, Charleston, IL, USA
jhmace@eiu.edu
<https://www.eiu.edu/psych/>

doi:10.1017/S0140525X23000213, e368

Abstract

This commentary supports Barzykowski and Moulin’s model, but departs from it on the question of functionality, where IAMs and déjà vu fractionate. The authors seem to say that IAMs are functional, while déjà vu is not. As there is no hard evidence supporting the idea that IAMs are functional, I argue that both phenomena should be viewed as cognitive failures.

Barzykowski and Moulin (B&M) present an interesting model for understanding involuntary autobiographical memories (IAMs) and déjà vu, two mysterious cognitive phenomena. They argue that déjà vu is the likely non-functional by-product (or “side effect”) of other cognitive processes, and while they make similar arguments for IAMs, they seem to indicate that IAMs have, in contrast, many functions (see Table 1 in their article). It is at this point, the question of adaptive functionality (e.g., Baddeley, 1988; Bluck & Alea, 2002), that I depart from their model. It is not clear to me how one (déjà vu) is seen as a cognitive failure and the other (IAM) is not. There is plenty of evidence that IAMs are by-products of other processes (e.g., unique cuing and priming; Ball, 2015; Berntsen, Staugaard, & Sørensen, 2013; Johannessen & Berntsen, 2010; Mace, 2005; Mace & Kruchten, 2022; Mace, McQueen, Hayslett, Staley, & Welch, 2019; Mace & Unlu, 2020, see target article). It is also reasonable to assume, even without this evidence, that IAMs, like déjà vu, may merely be cognitive failures. Thus, my commentary focuses on the idea that IAMs, like déjà vu and other similar phenomena, may be cognitive failures.

Proponents of the view that IAMs are functional have a number of challenges to overcome. For example, questionnaire studies on

IAM functions report functional hit rates ranging from one- to two-thirds (e.g., Mace & Atkinson, 2009; Rasmussen & Berntsen, 2011; Rasmussen, Ramsgaard, & Berntsen, 2015). One challenge is to explain how an evolved adaptive mechanism functions with such a low hit rate, presuming these numbers can be accepted. Relatedly, many of the reported functions in these studies can be seen as dubious (e.g., an IAM was entertaining or it made one feel better). Function theorists will also need to show how responses like these are functional, and not merely the function of the means used to assess the question, the survey method. Another challenge is to explain how an evolved adaptive mechanism, one which presumably reads the environment in some way to produce needed memories, interfaces with the cuing and priming variables to accomplish the same task. Indeed, aligning the function hypothesis with variables that seem to independently produce IAMs may prove challenging.

One additional explanatory variable for IAMs is the concept of retrieval mode (Tulving, 1983), a phenomenon which the target article addresses. Though originally discussed by Tulving (1983) as a tentative proposition, there is now good evidence that when the cognitive system is focused on intentional retrieval that stimuli are treated as episodic cues and memories of the past come to mind automatically and effortlessly (e.g., Herron, 2018; see also Rugg & Wilding, 2000). This process has been shown to be a routine, and often dominant, part of voluntary recall in autobiographical memory (e.g., Uzer, Lee, & Brown, 2012, known as direct retrieval, see also Conway, 2005, and the literature on direct retrieval). It is certainly possible that some IAMs may be the product of retrieval mode, a state which may be initiated inadvertently by aspects of the internal and external environment (e.g., Mace, 2010). If some IAMs are the products of retrieval mode gone awry, then this variable, too, would need to be aligned with the functional mechanism approach. However, this could be the more promising case if one could show that all IAMs were a function of the retrieval mode state, and such a state was not functioning erratically, but purposefully. This would raise a set of additional questions (e.g., the role of intent, or covert intent), and it clearly would be difficult to establish (but see Herron, 2018, and the retrieval mode and orientation literature).

Another problem concerns individual perceptions of spontaneous processes. Perceptions of IAMs are likely to differ greatly from déjà vu and other spontaneous processes, like action slips (e.g., Norman & Shallice, 1986; Reason, 1979). A déjà vu is likely to be seen as a false impression, an impossibility, and the circumstance may often reveal that. The experience of IAMs, though, can be quite different. They can be perceived as pleasant, circumstantially congruent relivings of one's past (e.g., Berntsen, 1998). As snippets of one's past, individuals may be more likely to imbue them with meaning, perceive them as functional, or imagine that they must be, even if they cannot see how. Such perceptions are likely to taint any self-report measure, and the design of such measures may only reinforce them. Given that experimental means are currently unavailable to answer questions of function (e.g., see interesting discussions in Downes, 2015), researchers have been forced to rely on self-report measures, like surveys. This presents another challenge for function theorists, as they also need to show that the data from such approaches are largely free of bias, or they can distinguish biased from unbiased data. A good example of potentially biased data are the dubious responses mentioned above, which do account for a substantial portion of the data (see Rasmussen et al., 2015). Researchers need to be cognizant of how personal biases might affect survey data, and guard against their own.

Thus, there are many challenges to the position that IAMs are functional, or sometimes can be. Because reliable means do not exist to adequately answer this question, the default, interim position, that IAMs are cognitive failures like déjà vu, appears to be the logical and prudent choice. Involuntary remembering can be explained in absence of a functional account. A number of variables have been identified as sources of IAMs (e.g., cuing, priming, and perhaps retrieval mode), variables which appear to independently cause them. B&M's model of IAM and déjà vu may work better if they presume the default position for both phenomena. Their model, and others, should treat IAM and déjà vu similarly if and until the function question can be answered. Alternative use accounts are possible (i.e., how memories might get used, Bluck, Alea, Habermas, & Rubin, 2005), but they should not be taken for function accounts.

Financial support. The writing of this commentary was not funded.


Competing interest. None.

References

- Baddeley, A. (1988). But what the hell is it for? In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues, vol. 1. Memory in everyday life* (pp. 3–18). John Wiley & Sons.
- Ball, C. T. (2015). IAMs and restrained eating. *Consciousness and Cognition*, 33, 237–244.
- Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, 6, 113–141.
- Berntsen, D., Staugaard, S. R., & Sørensen, L. M. T. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General*, 142, 426–444.
- Bluck, S., & Alea, N. (2002). Exploring the functions of autobiographical memory: Why do I remember the autumn? In J. D. Webster & B. K. Haight (Eds.), *Critical advances in reminiscence work: From theory to application* (pp. 61–75). Springer Publishing Company.
- Bluck, S., Alea, N., Habermas, T., & Rubin, D. C. (2005). A TALE of three functions: The self-reported uses of autobiographical memory. *Social Cognition*, 23, 97–117.
- Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, 53, 594–628.
- Downes, S. M. (2015). Evolutionary psychology, adaptation and design. In T. Heams, P. Huneman, G. Lecointre & M. Silberstein (Eds.), *Handbook of evolutionary thinking in the sciences* (pp. 659–673). Springer.
- Herron, J. E. (2018). Direct electrophysiological evidence for the maintenance of retrieval orientations and the role of cognitive control. *NeuroImage*, 172, 228–238.
- Johannessen, K. B., & Berntsen, D. (2010). Current concerns in involuntary and voluntary autobiographical memories. *Consciousness and Cognition*, 19, 847–860.
- Mace, J. H. (2005). Priming involuntary autobiographical memories. *Memory*, 13, 874–884.
- Mace, J. H. (2010). Involuntary remembering and voluntary remembering: How different are they? In J. H. Mace (Ed.), *The act of remembering: Toward an understanding of how we recall the past* (pp. 183–201). Wiley-Blackwell.
- Mace, J. H., & Atkinson, E. (2009). Can we determine the functions of everyday involuntary autobiographical memories? In M. Kelly (Ed.), *Applied memory* (pp. 199–212). Nova Science Publishers.
- Mace, J. H., & Kruchten, E. A. (2022). Semantic-to-autobiographical memory priming causes involuntary autobiographical memory production: The effects of single and multiple prime presentations. *Memory & Cognition*, 51, 115–128.
- Mace, J. H., McQueen, M. L., Hayslett, K. E., Staley, B. A., & Welch, T. J. (2019). Semantic memories prime autobiographical memories: Implications for everyday autobiographical remembering. *Memory & Cognition*, 47, 299–312.
- Mace, J. H., & Unlu, M. (2020). Semantic-to-autobiographical memory priming occurs across multiple sources: Implications for autobiographical remembering. *Memory & Cognition*, 48, 931–941.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz & D. E. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 1–14). Plenum Press.
- Rasmussen, A. S., & Berntsen, D. (2011). The unpredictable past: Spontaneously retrieved autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition*, 20, 1842–1846.
- Rasmussen, A. S., Ramsgaard, S. B., & Berntsen, D. (2015). The frequency and functions of involuntary and voluntary autobiographical memories across the day. *The Psychology of Consciousness: Theory, Research and Practice*, 2, 185–205.
- Reason, J. (1979). Actions not as planned. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness* (pp. 67–89). Academic Press.

- Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends Cognitive Sciences*, 4, 108–115.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press.
- Uzer, T., Lee, P.J., & Brown, N.R. (2012). On the prevalence of directly retrieved autobiographical memories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 38, 1296–1308.

The need for a unified framework: How Tulving's framework of memory systems, memory processes, and the SPI-model can guide and sharpen the understanding of déjà vu and involuntary autobiographical memories and add to conceptual clarity

Hans J. Markowitsch^a , Andreas Kordon^{b,c}
and Angelica Staniloiu^{a,b}

^aPhysiological Psychology, University of Bielefeld, Bielefeld, Germany;
^bSchwarzwaldklinik Hornberg, Hornberg, Germany and ^cDepartment of
Psychiatry, University of Freiburg, Germany
hjmarkowitsch@uni-bielefeld.de
andreas.wahl-kordon@oberbergkliniken.de
astaniloiu@uni-bielefeld.de

doi:10.1017/S0140525X23000055, e369

Abstract

Barzykowski and Moulin link déjà vu and involuntary autobiographical memories to the process of retrieval. They make no reference to Tulving's SPI-model. In this, it is proposed that information is acquired serially (S), stored in parallel (P), and retrieved independently (I). This model offers an alternative, elegant, view of involuntary autobiographical memory retrieval, as well as of déjà vu.

Memory entails encoding, consolidating, storage, and retrieving information. These are named processing stages of information and they rely on complex networks in the brain (Tulving & Markowitsch, 1998). Distinct from *memory processes* are the *memory systems* which have been proposed since the 1970s (e.g., Mishkin & Petri, 1984; Squire, 2004; Tulving, 1972) and refined over time (e.g., Schacter & Tulving, 1994; Schacter, Wagner, & Buckner, 2000; Staniloiu, Kordon, & Markowitsch, 2020). A sketch of memory systems is given in Figure 1.

Long-term memory systems have been partitioned into episodic-autobiographical memory (memory for personal events or experiences), semantic memory (conscious knowledge of facts, including factual self-knowledge), perceptual memory (conscious familiarity judgments), procedural memory (mechanical, motor-related skills), and priming memory (higher likelihood of re-identifying previously perceived stimuli) systems.

The “perceptual memory system” was identified and described as a legitimate distinct long-term memory system later than the other

four long-term memory (Tulving & Markowitsch, 1998). In contrast to the priming and procedural memory systems, this system acts “consciously” (noetically), but on a presemantic level and relies on familiarity judgments. An example is the conscious (noetic) identification of an apple without hesitation, no matter what color it has or whether it is already half eaten or not. Patients with semantic dementia may therefore still be able to distinguish, for example, an apple from a peach or pear by accessing perceptual representations of information via the perceptual memory system (despite losing capabilities for language and semantic memory).

In 1995, Tulving proposed his SPI-model which states that encoding of information follows a regular sequence – that means it is serial in that way that first simple, implicitly functioning, memory systems are engaged and only at the end of the series explicit, episodic encoding occurs. Information then is stored in parallel memory systems in the brain, and, most importantly it can be retrieved independently from the systems used for the encoding process (SPI = Serial, Parallel, Independent) (Tulving & Markowitsch, 1998, Fig. 2).

Seen in the light of the partitioning given in Figure 1, “involuntary autobiographical memories,” as described in the target article, belong to the *episodic-autobiographical memory system* (EAM), as they are, according to the authors, “recollections of the personal past”; on the other hand, they are retrieved automatically (“spontaneously,” as the authors write), without conscious effort, and therefore we argue that the retrieval of such personal past information recruits additional implicit/priming memory system resources (*priming system*) (Dew & Cabeza, 2011; Fehr, Staniloiu, Markowitsch, Erhard, & Herrmann, 2018; Henke, 2010). This recruitment of the priming memory resources may be facilitated at the neural level by a certain common neural mechanism that underpins implicit and explicit retrieval (Henke, 2010).

“Déjà vu” are described as “brief experiences of familiarity.” In another paper, one of the present authors writes that for déjà vu “the experi[m]ent is aware of a conflict in mental evaluations” (O'Connor, Wells, & Moulin, 2021, p. 835). We propose that in déjà vu, the conscious experience of familiarity that arises in the absence of corroboration from the semantic or episodic-autobiographical systems (recollection-familiarity) is due to a faulty/false retrieval/recognition in the conscious (noetic) perceptual memory system plus or minus priming system. Along this line, we provided evidence that false memories (false recognitions) do not only occur in semantic and EAM systems, but may also occur in other memory systems, such as the procedural memory system (see Borsutzky, Fujiwara, Brand, & Markowitsch, 2010).

We furthermore propose that information that is independently retrieved from the semantic or EAM systems according to the SPI-model along with sufficient intact metacognitive processes involved in monitoring and controlling retrieval (Risius et al., 2013) can decrease the strength of the feeling of familiarity, the confidence in the perceptual memory-based recognition judgments, and favor the interpretation that the conscious familiarity experience (false recognition) in déjà vus is false or impossible.

In line with the SPI-model, we (Staniloiu & Markowitsch, 2012) have pointed to the potential relationship between familiarity-based recognition, which perceptual memory (noetic, conscious) is intimately linked to, and familiarity linked to anoetic memory systems (priming, unconscious). Hereby, we argue that, reflecting the “porous boundaries” or dynamic exchange that exists between diverse memory systems at both the behavioral and neural levels (Dew & Cabeza, 2011), the SPI-model offers support for the view that the retrieval of IAMs and déjà vus engages an interplay

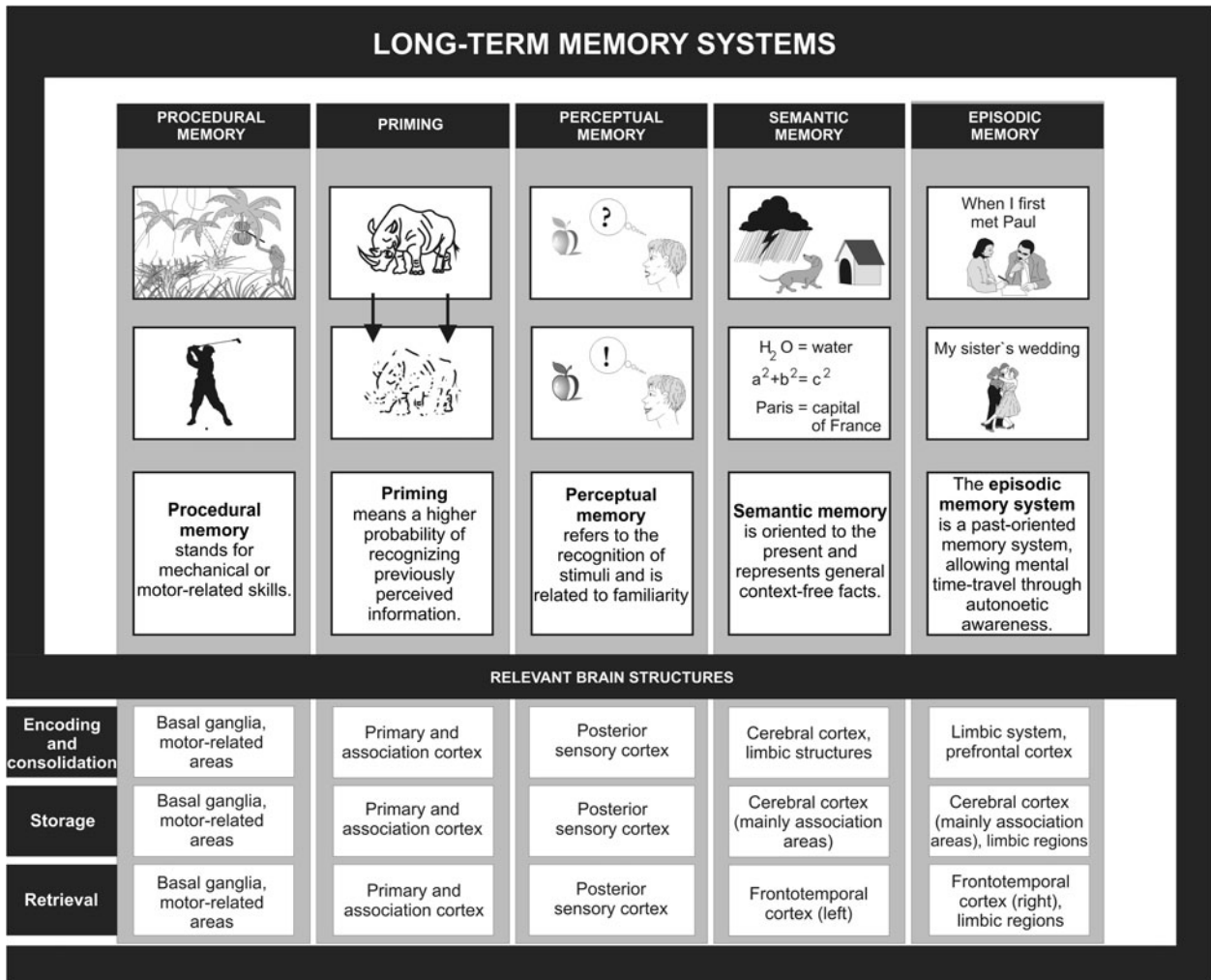


Figure 1 (Markowitsch et al.). The five long-term memory systems. *Procedural memory* is principally motor-based, but includes also sensory and cognitive skills and routines. *Priming* refers to a higher likelihood of re-identifying previously perceived information. *Perceptual memory* allows distinguishing and retrieving an item, an object, or a person based on distinct features. *Semantic memory* is context-free and refers to general facts; it encompasses general knowledge of the world. The *episodic-autobiographical memory (EAM) system* is context-specific with respect to time and place. It allows mental time travel and is based on self-reflection (autothesis). Examples are events such as a meeting with friends last week or the last celebration of New Year. The five systems develop phylogenetically and ontogenetically from left to right. With respect to dimensions of consciousness, Tulving (2005) considered the first two memory systems as being *anoetic*, the next two as *noetic*, and the EAM to be *autothetic* (“self-conscious”). A version of this sketch was created together with Endel Tulving.

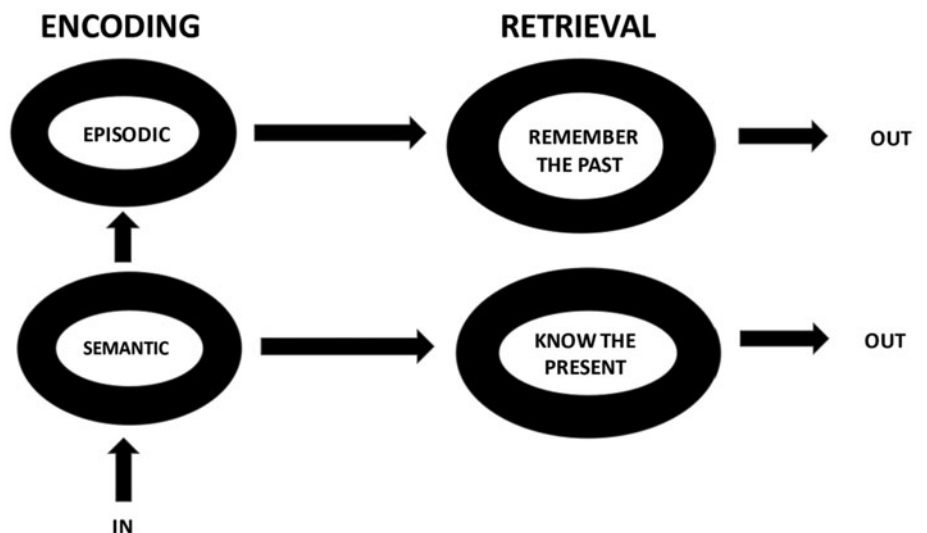


Figure 2 (Markowitsch et al.). Tulving's *SPI-model* (Tulving, 1995; Tulving & Markowitsch, 1998), exemplified for the *semantic* and *episodic* (or *episodic-autobiographical*) memory systems. Information can be encoded into semantic memory independently of episodic (-autobiographical) memory, but must be encoded into episodic(-autobiographical) memory “through” semantic memory (This was proven as well for ontogenetic learning of children; see Nelson & Fivush, 2004.) Encoded and stored information is potentially available for retrieval from one of the two systems, or from both of them (or – in generalizing – from all five of them, mentioned in Fig. 1). (Composed after Fig. 1 from Tulving & Markowitsch, 1998.)

between the (anoetic) priming memory system and auto-noetic and noetic memory systems (see Fig. 2).

Financial support. There was no funding of this contribution.

Competing interest. None.

References

- Borsutzky, S., Fujiwara, E., Brand, M., & Markowitsch, H. J. (2010). Susceptibility to false memories in patients with ACoA aneurysm. *Neuropsychologia*, *48*, 2811–2823.
- Dew, I. T. Z., & Cabeza, R. (2011). The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. *Annals of the New York Academy of Sciences*, *1224*, 174–190.
- Fehr, T., Staniloiu, A., Markowitsch, H. J., Erhard, P., & Herrmann, M. (2018). Neural correlates of free recall of “famous faces” in a “hypermnestic” individual as compared to an age- and education-matched reference group. *BMC Neuroscience*, *19*, Art. 35, 1–12. <https://doi.org/10.1186/s12868-018-0435-y>
- Henke, K. A. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, *11*, 523–532.
- Mishkin, M., & Petri, H. L. (1984). Memories and habits: Some implications for the analysis of learning and retention. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 287–296). Guilford Press.
- Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, *111*, 486–511.
- O'Connor, A. R., Wells, C., & Moulin, C. J. A. (2021). Déjà vu and other dissociative states in memory. *Memory*, *29*, 835–842.
- Risius, U.-M., Staniloiu, A., Piefke, M., Maderwald, S., Schulte, F., Brand, M., & Markowitsch, H. J. (2013). Retrieval, monitoring and control processes: A 7 Tesla fMRI approach to memory accuracy. *Frontiers in Behavioral Neuroscience*, *7*, Art. 24, 1–21. <https://doi.org/10.3389/fnbeh.2013.00024>
- Schacter, D., & Tulving, E. (1994). What are the memory systems of 1994? In D. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 1–38). MIT Press.
- Schacter, D. L., Wagner, A. D., & Buckner, R. L. (2000). Memory systems of 1999. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 627–644). Oxford University Press.
- Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*, 171–177.
- Staniloiu, A., Kordon, A., & Markowitsch, H. J. (2020). Quo vadis episodic memory – past, present, perspective. *Neuropsychologia*, *141*, Art. 10736, 1–8. <https://doi.org/10.1016/j.neuropsychologia.2020.107362>
- Staniloiu, A., & Markowitsch, H. J. (2012). The remains of the day in dissociative amnesia. *Brain Sciences*, *2*, 101–129.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). Academic Press.
- Tulving, E. (1995). Organization of memory: Quo vadis? In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 839–847). MIT Press.
- Tulving, E. (2005). Episodic memory and auto-noesis: Uniquely human? In H. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Evolution of self-knowing consciousness* (pp. 3–56). Oxford University Press.
- Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus*, *8*, 198–204.

On the frequency and nature of the cues that elicit déjà vu and involuntary autobiographical memories

Ricardo Morales-Torres and Felipe De Brigard 

Department of Psychology and Neuroscience, and Center for Cognitive Neuroscience, Duke Institute for Brain Sciences, Duke University, Durham, NC, USA
ricardo.morales.torres@duke.edu
felipe.debrigard@duke.edu
<https://www.imclab.org/people>

doi:10.1017/S0140525X23000134, e370

Abstract

Barzykowski and Moulin suggest that déjà vu and involuntary autobiographical memories recruit similar retrieval processes. Here, we invite the authors to clarify three issues: (1) What mechanism prevents déjà vu to happen more frequently? (2) What is the role of semantic cues in involuntary autobiographical retrieval? and (3) How déjà vu relates to non-believed memories?

Barzykowski and Moulin (B&M henceforth) offer an intriguing model to unify déjà vu and involuntary autobiographical memories (IAM). A critical claim of their model is that both psychological phenomena share, as a common cause, the automatic processing of ongoing cues in the environment. Specifically, they suggest that a perceptual overlap between an environmental cue and a particular mnemonic content is the starting point for both déjà vu and IAM. While their proposal is compelling, our comment raises three points that, we think, deserve further elaboration. The first and second points concern the nature of the cues leading to déjà vu and IAM experiences. The third point draws attention to the mechanism that differentiates déjà vu from non-believed memories.

The first concern can be put simply: There seems to be a tension between the fact that perceptual overlap is an extremely common phenomenon whereas déjà vu, as remarked by B&M, is extremely uncommon. How come such a rare psychological phenomenon is largely explained by such a common occurrence as the overlap between a memory representation and an environmental cue? According to B&M's model of déjà vu, our memory system involves a continuously active mechanism constantly scanning the environment for cues that could match memory representations. When this mechanism finds a cue that matches a stored representation, but the activation is not sufficient to elicit autobiographical recall, people experience only a sense of familiarity. This feeling is then metacognitively assessed in order to evaluate its source. If the system fails to find the source of the feeling of familiarity, it evaluates the cue as novel, and the output is thus experienced as déjà vu.

The problem is that the model is not sufficiently clear as to how exactly this underlying mechanism manages to discriminate between the very frequently experienced environmental cues that overlap with stored memory representations and that never get experienced as déjà vu, and the very infrequent instances in which they do. To be sure, we agree with B&M in that our memory system is constantly presented with overlaps between environmental cues and mnemonic representations (Fernández & Morris, 2018; Kafkas & Montaldi, 2018). What we find unclear is how and why exactly some of those overlaps manage to only elicit a sense of familiarity and a metacognitive evaluation of novelty whereas the vast majority elicit no feeling at all. One possible avenue to fill this gap in the model could be found in the schema literature, whereby novelty is proportional to the mismatch between the stimulus and its context (Van Kesteren, Ruitter, Fernández, & Henson, 2012). Another possible avenue is to appeal to prediction error, whereby novelty relates to the divergence between a prior likelihood and a posterior probability (De Brigard, 2012; Ergo, De Loof, & Verguts, 2020). Either way, we believe this bit of the model needs further development.

The second concern is about the relationship between semantic memory and IAM. When talking about IAM, the authors

swiftly mention that semantic information can elicit IAM. Indeed, there is plenty of evidence suggesting that abstract and conceptual information are powerful triggers of IAM (Mace, 2005; Mace & Hidalgo, 2022). However, almost all of the examples B&M use to substantiate the model involve IAM that are triggered by perceptual cues. While we certainly agree that sensory cues are an important and frequent environmental trigger for IAM (Conway & Pleydell-Pearce, 2000), we also believe that B&M's model could be strengthened by clarifying the distinct contribution of the perceptual and the semantic properties of the environmental cues that can trigger IAM. Consider how semantic information can trigger IAMs in semantic-to-autobiographical memory priming. The work of Mace, McQueen, Hayslett, Staley, and Welch (2019) has shown that priming a concept (e.g., reading the word "frog") increases the likelihood of having an IAM related to that concept. Moreover, perceptual cues contain both sensory and semantic information that can influence episodic memory (Davis et al., 2021; Hovhannisyian et al., 2021). Each of these could in turn trigger IAM in different ways. For example, when seeing a Darwin's frog, its sensory properties (such as its form or color) could trigger an IAM about a pet frog in my childhood, whereas the conceptual information associated with seeing such a frog could generate an IAM about the period of my life in which I studied the theory of evolution. By further exploring the properties of the cues that trigger IAM, B&M could generate a more thorough explanation about the memory processes that support IAM.

Our final point is a question about the relationship between déjà vu and another surprising memory phenomenon: "Non-believed memories." In their paper, B&M state that déjà vu "arises out of a higher order interpretation of retrieval processes; it is not possible to have a déjà vu experience and not be aware of it. This critical feature distinguishes it from a false memory." It is tempting to think that a similar metacognitive evaluation helps to explain "non-believed memories," that is, recollective experiences of events one no longer believes that happened (Otgaar, Scoboria, & Mazzoni, 2014; Scoboria, Boucher, & Mazzoni, 2015). While unusual, non-believed memories have been shown to occur in around 20% of the population, and present similar phenomenological features as real autobiographical memories (Scoboria, Mazzoni, Kirsch, & Relyea, 2004). Moreover, like déjà vu, non-believed memories also involve the metacognitive awareness that what is experienced is not an actual memory. What is surprising, though, is that while in the case of déjà vu the metacognitive judgment that the event is not a memory is sufficient to remove its recollective experience, the same is not the case with non-believed memories: People know the events they thought were memories are not, but still experience them as such. What accounts for this difference? We think it would be interesting to try to expand the proposed model to see how it could accommodate non-believed memories.

Financial support. None.


Competing interest. None.

References

- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107(2), 261.
- Davis, S. W., Geib, B. R., Wing, E. A., Wang, W.-C., Hovhannisyian, M., Monge, Z. A., & Cabeza, R. (2021). Visual and semantic representations predict subsequent memory in perceptual and conceptual memory tests. *Cerebral Cortex*, 31(2), 974–992.

- De Brigard, F. (2012). Predictive memory and the surprising gap. Commentary on Andy Clark's "Whatever next? Predictive brains, situated agents and the future of cognitive science". *Frontiers in Psychology: Theoretical and Philosophical Psychology*, 3, 420.
- Ergo, K., De Loof, E., & Verguts, T. (2020). Reward prediction error and declarative memory. *Trends in Cognitive Sciences*, 24(5), 388–397.
- Fernández, G., & Morris, R. G. M. (2018). Memory, novelty and prior knowledge. *Trends in Neurosciences*, 41(10), 654–659.
- Hovhannisyian, M., Clarke, A., Geib, B. R., Cicchinelli, R., Monge, Z., Worth, T., ... Davis, S. W. (2021). The visual and semantic features that predict object memory: Concept property norms for 1,000 object images. *Memory & Cognition*, 49(4), 712–731.
- Kafkas, A., & Montaldi, D. (2018). How do memory systems detect and respond to novelty? *Neuroscience Letters*, 680, 60–68.
- Mace, J. H. (2005). Priming involuntary autobiographical memories. *Memory*, 13(8), 874–884. <https://doi.org/10.1080/09658210444000485>
- Mace, J. H., & Hidalgo, A. M. (2022). Semantic-to-autobiographical memory priming affects involuntary autobiographical memory production after a long delay. *Consciousness and Cognition*, 104, 103385. <https://doi.org/10.1016/j.concog.2022.103385>
- Mace, J. H., McQueen, M. L., Hayslett, K. E., Staley, B. J. A., & Welch, T. J. (2019). Semantic memories prime autobiographical memories: General implications and implications for everyday autobiographical remembering. *Memory & Cognition*, 47(2), 299–312.
- Otgaar, H., Scoboria, A., & Mazzoni, G. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science*, 23(5), 349–354. <https://doi.org/10.1177/0963721414542102>
- Scoboria, A., Boucher, C., & Mazzoni, G. (2015). Reasons for withdrawing belief in vivid autobiographical memories. *Memory*, 23(4), 545–562. <https://doi.org/10.1080/09658211.2014.910530>
- Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology*, 18(7), 791–807.
- Van Kesteren, M. T. R., Ruiters, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, 35(4), 211–219.

Déjà vu may be illusory gist identification

Shen Pan  and Peter Carruthers

Department of Philosophy, University of Maryland, College Park, MD, USA
shenpan@umd.edu
pcarruth@umd.edu
<https://sites.google.com/site/theshenpan/>
<https://faculty.philosophy.umd.edu/pcarruthers/>

doi:10.1017/S0140525X23000195, e371

In déjà vu, a novel experience feels strangely familiar. Here we propose that this phenomenology is best seen as consisting of an illusory feeling of identification of the gist of the current scene or event, rather than in the intensity of the fluency-based, metacognitive feeling of familiarity.

Barzykowski and Moulin (B&M) propose that both involuntary autobiographical memories (IAMs) and déjà vu result from the same memory retrieval processes, initiated by "a continuously active memory system that automatically and rapidly scans the environment for matching representations." We are largely in agreement with their overall framework. Here we focus on B&M's discussion of déjà vu, which they characterize in terms of a false familiarity devoid of content (thus unlike IAMs, which do deliver content). B&M emphasize that this feeling arises when retrieval fails, contending that the relevant notion of familiarity is a uniquely phenomenological one (hereafter, "familiarity_p"). Notably, this is a slightly different notion of familiarity from that employed in the recognition-memory literature (hereafter, "familiarity_r"), one that refers to a decision-making process based on feelings of fluency (Yonelinas, 2002). But

B&M's view is that the feeling of familiarity_p, characteristic of déjà vu results from a mismatch between familiarity_f and one's expectations about fluency. Thus, the false familiarity_p constitutes a higher-order, metacognitive interpretation of retrieval processes, and on this view, familiarity_p naturally admits of degrees of intensity, corresponding to greater or lesser mismatch.

B&M are far from alone in adopting a familiarity_f-based conception of déjà vu. It makes intuitive sense that in having a déjà vu experience, one feels the current scene or event to be strangely familiar. We emphasize, however, that this intuitive, colloquial sense of familiarity is not fully captured by either familiarity_f or familiarity_p. In déjà vu, one does not merely experience phenomenological fluency; nor is it just the feeling that the felt fluency is unexpected. Adding intensity does not help, either. Uniquely characteristic of déjà vu is the impression that *this exact scene or event has happened before*, a feeling which one simultaneously judges to be implausible (since, e.g., one is visiting a city for the first time). The phenomenology of déjà vu thus has a specificity component, despite the fact that no contextual specifics are successfully retrieved. But this is precisely why déjà vu is so striking, and it suggests that the sense of familiarity of relevance to déjà vu is something more akin to an illusory feeling of identification (cf. Cleary, 2008).

Let us put the point differently. Note that recollection and familiarity – the two bases of recognition – correspond to “remembering” and “knowing,” respectively. Familiarity_f-based accounts of déjà vu thus characterize the phenomenology in terms of erroneously *knowing* that something has been encountered before. Phenomenologically, this characterization is inadequate. Indeed, both of the two déjà vu quotes with which B&M begin their article express an illusory feeling of identifying the current scene or event with the personal past (e.g., “Each time I feel a strong conviction that I’ve seen all of it happen before”). To be sure, this is not full-fledged recollection/remembering; but it is not merely familiarity/knowing, either. Since B&M seem friendly to the idea of familiarity and recollection lying along a continuum, we suggest that identification may lie somewhere in between.

We have argued that the phenomenology of déjà vu is best seen as an illusory feeling of identification, rather than a mere feeling of familiarity. We now argue that the feeling of identification is not fully devoid of content. Rather, it can plausibly be explained as recollection-based identification of the *gist* of the current scene or event, when that gist matches one (and only one) that is stored in memory.

It is now well known that ensemble perception (or “gist perception”) takes place alongside focused object perception, and is responsible for much of the sense of richness that attaches to conscious experience (Cohen, Dennett, & Kanwisher, 2016; McClelland & Bayne, 2016; Whitney & Leib, 2018). Thus, perceivers can accurately extract the average orientation of a set of lines, or the average size of a set of circles, or the average expression on a set of faces; and they can do so quite swiftly, often within 50 milliseconds, or (in cases of sequential presentation) at presentation speeds of 20 images per second (Chong & Treisman, 2003; Haberman & Whitney, 2009). Moreover, they can extract a number of distinct kinds of gist at once, such as the average emotional expression, the average speed, and the average direction of movement of a crowd of people (Haberman, Lee, & Whitney, 2015; Sweeny, Haroz, & Whitney, 2013; Sweeny & Whitney, 2014), although accuracy on each individual dimension tends to drop

as more and more dimensions get added (Emmanouil & Treisman, 2008).

It is also known that the gist properties of an event or scene get stored in long-term memory alongside episodic details (Brady & Alvarez, 2011; Lew & Vul, 2015). When an episodic autobiographical memory is accessed or triggered and becomes conscious, then, it will characteristically comprise both episodic details (who was there, what happened and when), together with background gist – the scene was a picnic in a park, with other people distributed against a background of grass and distant trees; or the scene was a busy street with tall buildings and traffic moving slowly past.

B&M speculate that perceived scenes may be frequent triggering cues for both IAMs and déjà vu. Whereas they see this as another possible commonality between IAMs and déjà vu, it might also suggest a way in which the two are interestingly different. Scenes tend to be rich in local features, thus providing numerous sources of conceptual or perceptual overlap with stored memory representations, activation of sufficient numbers of which will be apt to pass the threshold for entry into consciousness, generating an IAM. But scene perception also involves rapid processing of global information, specifically gist information, as we have noted. This will then either be matched with many stored representations (“cue overload,” in which case none pass the threshold for consciousness individually, unless tied to an IAM), or it may – unexpectedly – match just one without activating sufficient detail to create an IAM. In the latter case, the result, we suggest, is gist identification combined with knowledge that the situation is novel (a déjà vu experience).


Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

- Brady, T., & Alvarez, G. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22, 384–392.
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, 43, 393–404.
- Cleary, A. M. (2008). Recognition memory, familiarity, and déjà vu experiences. *Current Directions in Psychological Science*, 17(5), 353–357.
- Cohen, M. A., Dennett, D., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences*, 20, 324–335.
- Emmanouil, T., & Treisman, A. (2008). Dividing attention across feature dimensions in statistical processing of perceptual groups. *Perception and Psychophysics*, 70, 946–954.
- Haberman, J., Lee, P., & Whitney, D. (2015). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *Journal of Vision*, 15(4), 1–11.
- Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 718–734.
- Lew, T., & Vul, E. (2015). Ensemble clustering in visual working memory biases location memories and reduces the Weber noise of relative positions. *Journal of Vision*, 15(4), 10.
- McClelland, T., & Bayne, T. (2016). Ensemble coding and two conceptions of perceptual sparsity. *Trends in Cognitive Sciences*, 20, 641–642.
- Sweeny, T., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 329–337.
- Sweeny, T., & Whitney, D. (2014). Perceiving crowd attention: Ensemble perception of a crowd's gaze. *Psychological Science*, 25, 1903–1913.
- Whitney, D., & Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, 69, 12.1–12.25.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46, 441–517.

Accommodating the continuum hypothesis with the déjà vu/déjà vécu distinction

Denis Perrin 

Centre for Philosophy of Memory/Institut de Philosophie de Grenoble, Univ. Grenoble Alpes, Grenoble, France

denis.perrin@univ-grenoble-alpes.fr

<https://phil-mem.org/members/perrin.php>

doi:10.1017/S0140525X23000225, e372

Abstract

On Barzykowski and Moulin's continuum hypothesis, déjà vu and involuntary autobiographical memories (IAMs) share their underpinning neurocognitive processes. A discontinuity issue for them is that familiarity and episodic recollection exhibit different neurocognitive signatures. This issue can be overcome, I say, provided the authors are ready to distinguish a déjà vécu/episodic IAM continuity and a déjà vu/semantic IAM continuity.

On Barzykowski and Moulin's continuum hypothesis, the feeling of familiarity involved in déjà vu experiences and the episodic recollection involved in involuntary autobiographical memories (IAMs) are due to the same underpinning neurocognitive retrieval processes. Elaborating their hypothesis with a threshold model, the authors propose that cues in the environment constantly activate autobiographical memory. As a result, some memories cross the awareness threshold due to sufficient activation and give rise to complete retrieval. But it also happens that memories, due to weak activation for instance, give rise to incomplete retrieval, or that no specific memory crosses the threshold and no content is retrieved. In these cases, the cue is predicted to feel familiar while no complete recollection, or no recollection at all, occurs.

Though such a single-mechanism model can avail itself of some recent proposals along the same lines, as Bastin et al. (2019) note, the majority of the existing models suggest that there is fractionation of both the memory processes (Yonelinas, 2002) and the dedicated brain regions (Aggleton & Brown, 1999) by reference to familiarity and recollection, in particular in the medial temporal lobe, with parahippocampal regions dedicated to familiarity and the hippocampal structure dedicated to recollection. Important recent theorizing of episodic memory as a distinct system specifically involving mental time travel feelings (Barry & Maguire, 2019, for instance) lend further support to the same suggestion. In brief, the feeling of familiarity and episodic recollection exhibit neatly distinct neurocognitive signatures. If these in-the-majority models are right, the continuum hypothesis is in a bad position, since one of its main goals is to establish that déjà vu and IAMs can be grouped as two forms of involuntary cognition unified by the same underlying retrieval process. There is therefore a discontinuity issue for the continuum hypothesis.

I think that the hypothesis can overcome this issue. But to do so, it must make room for a distinction neglected by its current formulation, and be ready for ensuing refinements.

Though the authors make no room for this, two types of déjà vu experience must be distinguished, arguably. Subjects can undergo the experience of reliving in the present the very same *episode of experience* as one they have already lived before. The author's fourth initial example describes precisely this: "It feels like you're living an experience that you've already lived through." But subjects can also undergo déjà vu experiences illustrated by reports like the following: "Frequently I relate this to seeing people in the street. I will 'Recognise them,' rack my brain trying to remember where and expect them to greet me as we pass" (Illman, 2012). In such experiences, subjects have the feeling that a particular perceived *item* like a person – versus a whole episode of experience – has been already encountered during some *indeterminate, potentially multiple* past episodes of experience – versus a *specific* episode of experience – which refers to the feeling of familiarity on a standard characterization of it. Some have judged these differences sufficiently important to call the first type of experiences *déjà vécu* and distinguished it from *déjà vu*, arguing that while in déjà vu a *feeling of familiarity* occurs, in déjà vécu an *episodic-recollection-related feeling* occurs (O'Connor, Lever, & Moulin, 2010).

On my diagnosis, the discontinuity issue originates in the authors' attempt to account for the continuity of déjà vu experiences with episodic IAMs by considering déjà vu instead of considering déjà vécu. This diagnosis suggests a way out.

Though déjà vécu is akin to episodic IAM, they are distinct phenomena. Episodic IAM consists of an *actual* recollection of a past episode of experience. By contrast, déjà vécu consists of the awareness of the ongoing episode of experience as having been lived before with *no actual* recollection of any previous occurrence of it. Following a recent elaboration of this point, while episodic recollection brings into play an *autonoetic feeling*, déjà vécu brings into play an *episodic feeling of knowing* (EFOK), namely the feeling that one could remember other occurrences of the ongoing episode of experience despite not carrying out this memory task at the moment (Perrin, Moulin, & Sant'Anna, forthcoming). Interestingly, a prevailing account of EFOK says that shared processes underly both this feeling and actual episodic recollection. According to Souchay, Moulin, Clarys, Taconnat, and Isingrini (2007), EFOK results from the subpersonal and partial retrieval of contextual details, phenomenological feelings, and self-awareness similar to the ones that occur in actual episodic recollection. My proposal is, then, that déjà vécu experience is underpinned by the same retrieval process as episodic IAM, a process that would be assessed as malfunctioning when such an experience occurs.

What about déjà vu? It is worth recalling at this stage that autobiographical memory, IAM included, comprises two main forms (Conway, 2001). Though the authors talk mainly about the episodic form, autobiographical memory also possesses a semantic form. My proposal is, then, that while EFOK-based déjà vécu experiences share their underpinning neurocognitive processes with *episodic* IAMs, feeling-of-familiarity-based déjà vu experiences share their underpinning neurocognitive processes with *semantic* IAMs. In effect, it has since long been noted that familiarity and semantic memory are closely related. As one remembers that one has already encountered in the past an item, the item typically feels familiar, and Tulving (1985) even assimilates the noetic consciousness characteristic of semantic memory to familiarity.

If those proposals are correct, the continuity issue pointed out above can be put to rest and the continuum hypothesis reinstated. As I have argued, the authors' version of the hypothesis is threatened because it considers only familiarity-based déjà vu experiences and posits neurocognitive continuity between this type of experiences and episodic IAMs. Once the déjà vu–déjà vécu distinction and the episodic-semantic IAMs distinction are clear, one can posit continuity relationships at the right place, namely between déjà vu experiences and semantic IAMs, on the one hand, and between episodic IAMs and déjà vécu experiences, on the other hand.

Financial support. None.

Competing interest. None.

References

- Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences*, 22, 425–498.
- Barry, D. N., & Maguire, E. A. (2019). Remote memory and the hippocampus: A constructive critique. *Trends in Cognitive Sciences*, 23(2), 128–142. <https://doi.org/10.1016/j.tics.2018.11.005>
- Bastin, C., Besson, G., Simon, J., Delhay, E., Geurten, M., Willems, S., & Salmon, E. (2019). An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, 42(281), 1–60. doi:10.1017/S0140525X19000621
- Conway, M. (2001). Sensory-perceptual episodic memory and its context: Autobiographical memory. *Philosophical Transactions: Biological Sciences*, 356(1413), 1375–1384.
- Illman, N. A. (2012). *Recognition memory impairments in temporal lobe epilepsy: The contribution of recollection and metacognition*. PhD thesis, University of Leeds (unpublished).
- O'Connor, A. R., Lever, C., & Moulin, C. J. A. (2010). Novel insights into false recollection: A model of déjà vécu. *Cognitive Neuropsychiatry*, 15, 118–144.
- Perrin, D., Moulin, C., & Sant'Anna, A. (forthcoming). Déjà vécu is not déjà vu: An ability view. *Philosophical Psychology*. <https://doi.org/10.1080/09515089.2022.2161357>
- Souchay, C., Moulin, C. J. A., Clarys, D., Taconnat, L., & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: Evidence from feeling of knowing and recollection. *Consciousness and Cognition*, 16, 769–784.
- Tulving, E. (1985). How many memory systems are there? *American Psychologist*, 40, 385–398.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46, 441–517.

From jamais to déjà vu: The respective roles of semantic and episodic memory in novelty monitoring and involuntary memory retrieval

Louis Renoult^a  and J. Bruno Debruille^b 

^aSchool of Psychology, University of East Anglia, Norwich, UK and ^bDepartment of Neuroscience, Frank B. Common Pavilion, Douglas Hospital Research Center, McGill University, Montreal, QC, Canada

L.Renoult@uea.ac.uk

bruno.debruille@mcgill.ca

<https://research-portal.uea.ac.uk/en/persons/louis-renoult>

<https://www.mcgill.ca/psychiatry/bruno-debruille>

doi:10.1017/S0140525X23000158, e373

Abstract

Barzykowski and Moulin's model proposes that déjà vu and involuntary autobiographical memories are the result of a continuously active memory system that tracks the novelty of situations. Déjà vu would only have episodic content and concern interpretation of prior experiences. We argue that these aspects of the model would gain to be clarified and explored further and we suggest possible directions.

The model of Barzykowski and Moulin (B&M) proposes an interesting comparison of involuntary autobiographical memories and déjà vu experience, whereby déjà vu would be the product of an underdeveloped form of involuntary memory retrieval. The model includes the attractive notion that familiarity/novelty detection mechanisms might need to be permanently active, sometimes leading to déjà vu experience and to false detections. Moreover, it is proposed that déjà vu would only have episodic content and concern interpretation of prior experiences, rather than semantic knowledge. These interesting aspects of their proposal are not elaborated upon, unfortunately. However, they are consistent with the finding that semantic and episodic memory appear to have different sensitivity to novelty processing. Repeated semantic processing has often been associated with the phenomenon of semantic satiation, a temporary loss of meaning attached to stimuli. Explicit attention to meaning is necessary to counteract these effects (Renoult, Wang, Mortimer, & Debruille, 2012). Once an object has been identified and categorized, the conceptual system may thus no longer be oriented to this particular object (Murphy, 2002). As shown in another work by the authors (Moulin, Bell, Turunen, Baharin, & O'Connor, 2021), repeated semantic processing may actually lead to the opposite experience, the so-called, jamais-vu (never seen).

In contrast, episodic memory allows to encode unique events (even if it is just a repetition of the same item, Tulving, 1972, 1983; or an event similar to a previous experience, Rubin & Umanath, 2015), while related neural processes may create distinct representations even for seemingly similar events (Nadel & Moscovitch, 1997; Norman, 2010; O'Reilly, Bhattacharyya, Howard, & Ketz, 2014). Therefore, if déjà vu experiences depend on constantly active novelty detection mechanisms, episodic memory would thus be more likely to be associated with these experiences, consistent with B&M's current proposal. Nonetheless, a better characterisation of these novelty detection mechanisms would help to better understand how the experiences of involuntary autobiographical memories and déjà vu are generated.

As a matter of fact, the points made in the article that déjà vu would “concern interpretation of prior experiences” and would typically happen for “familiar places and people” suggest a role for semantic knowledge. Consistent with the proposal that episodic memory operations typically require semantic memory (Tulving, 2002), one may ask, for example, whether déjà vu could be related to the retrieval of schema knowledge about a similar past experience, which would act as a cue for a (failed) episodic retrieval attempt. In other words, it could be that it is this activation of semantic knowledge about similar prior experiences that would trigger a retrieval mode: A “tonically maintained state” needed for episodic retrieval (Rugg & Wilding, 2000; Tulving, 1983).

Accessing conceptual information can indeed prime the retrieval of specific memories (Mace, McQueen, Hayslett, Staley, & Welch, 2019; Reiser, Black, & Abelson, 1985) and semantic processing of cues was proposed to be essential for episodic retrieval (Tulving, 1983). However, in the case of déjà vu, this retrieval attempt may fail, or conflict with the awareness that the experience is novel. Presumably, the experience of jamais vu (Moulin et al., 2021; O'Connor, Wells, & Moulin, 2021) would similarly not be associated with the retrieval of specific memories. However, in that case, it would be due to disrupted semantic retrieval (semantic satiation), which may also deprive episodic memory from a relevant cue. Whether these proposals are valid or not, further exploration of the notion that déjà vu would only have episodic content appears important, as well as clarifying the role of interpretative process (and of semantic knowledge) in generating these experiences.

Finally, as noted by B&M, the role of personal semantics (e.g., the type of personal semantics involved and the timing of its involvement) would be worth exploring further in the context of evaluating how it may influence or prime, as discussed in the article, the content of retrieval. For instance, activating context-dependent types of personal semantics (e.g., I visit this café every Monday) might more likely lead to the retrieval of involuntary autobiographical memories. More abstracted form of personal semantics (e.g., I take milk in my coffee) would more likely be restricted to feeling of familiarity or déjà vu (Sheldon, Peters, & Renoult, 2020). This would be consistent with the findings of Berntsen, Staugaard, and Sørensen (2013) that involuntary episodic memories are retrieved more often in response to specific compared to more generic cues.

We thoroughly enjoyed the proposal of B&M and hope that further research, based on their model, will help clarify these issues.

Financial support. Louis Renoult is supported by grant MR/S011463/1 from the Medical Research Council (MRC).

Competing interest. None.

References

- Berntsen, D., Staugaard, S. R., & Sørensen, L. M. T. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General*, *142*(2), 426–444.
- Mace, J. H., McQueen, M. L., Hayslett, K. E., Staley, B. J. A., & Welch, T. J. (2019). Semantic memories prime autobiographical memories: General implications and implications for everyday autobiographical remembering. *Memory and Cognition*, *47*(2), 299–312.
- Moulin, C. J. A., Bell, N., Turunen, M., Baharin, A., & O'Connor, A. R. (2021). The induction of jamais vu in the laboratory: Word alienation and semantic satiation. *Memory*, *29*(7), 933–942.
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*, 217–227.
- Norman, K. A. (2010). How hippocampus and cortex contribute to recognition memory: Revisiting the complementary learning systems model. *Hippocampus*, *20*, 1217–1227.
- O'Connor, A. R., Wells, C., & Moulin, C. J. A. (2021). Déjà vu and other dissociative states in memory. *Memory*, *29*(7), 835–842.
- O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2014). Complementary learning systems. *Cognitive Science*, *38*, 1229–1248.
- Reiser, B. J., Black, J. B., & Abelson, R. P. (1985). Knowledge structures in the organization and retrieval of autobiographical memories. *Cognitive Psychology*, *17*(1), 89–137.
- Renoult, L., Wang, X., Mortimer, J., & Debrulle, J. B. (2012). Explicit semantic tasks are necessary to study semantic priming effects with high rates of repetition. *Clinical Neurophysiology*, *123*, 741–754.
- Rubin, D. C., & Umanath, S. (2015). Event memory: A theory of memory for laboratory, autobiographical, and fictional events. *Psychological Review*, *122*(1), 1–23.
- Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends Cognitive Science*, *4*, 108–115.
- Sheldon, S., Peters, S., & Renoult, L. (2020). Altering access to autobiographical episodes with prior semantic knowledge. *Consciousness and Cognition*, *86*, 103039.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). Academic Press.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press.
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25.

Cueing involuntary memory

Sarah Robins  and Maziyar Afifi

Department of Philosophy, University of Kansas, Lawrence, KS, USA

skrobins@ku.edu

sarahkrobins.com

maziyarafifi@ku.edu

doi:10.1017/S0140525X23000122, e374

Abstract

We raise two points about cues, which complicate Barzykowski and Moulin's attempt at a unified model of memory retrieval. First, cues operate differently in voluntary and involuntary contexts. Second, voluntary and involuntary memory can be interconnected, as in cases of chaining.

We applaud Barzykowski and Moulin's (B&M) advocacy for a comprehensive account of memory retrieval. As they acknowledge, voluntary memory retrieval has long served as the default. The key terms and processes of retrieval are characterized on the basis of their operation in voluntary contexts. The success of a unified account is thus dependent on how easily concepts that were developed to account for voluntary memory phenomena transfer to involuntary memory phenomena.

Here, we focus specifically on the concept of cues. All memory is cued, at least in the sense that memory is the recurrence, re-experiencing, or reconstructive representation of prior information, events, or activities. For an item retained in memory, the potential for reactivation is more or less constant. A memory coming to mind *now*, rather than at some other point in time (or not at all), is explained at least in part by some internal or external thing serving as a cue for its reactivation.

As causes of memory retrieval, cues precede their effects. B&M's comprehensive retrieval model is built around this, using cues as the starting point for all retrieval processing (as depicted in their Fig. 1). While this is true of both voluntary and involuntary memory – that is, that they both start from cues – the details of this process are importantly different across these two forms of memory.

In voluntary remembering, there is a clearly identified *cue* which is used to prompt a clearly identified *target*. Within such a framework, there is a clear starting point and a clear end, as well as success conditions, making it possible to measure the retrieval speed, perceived fluency, and accuracy/error of the remembering attempt. Importantly, in such cases, the cue can be identified prior to and independently of the memory it activates. The causal process from the cue to retrieval is thus extended and consciously available to the would-be rememberer.

The same is not true for cases of involuntary remembering. Involuntary memories and déjà vu are cued, in the sense that

all memories are, but it is not possible to identify the cue prior to the involuntary activation of a memory. B&M characterize IAMs and déjà vu as cases of “spontaneous processing of cues.” This is, in effect, an acknowledgement that, in cases of involuntary retrieval, a cue can only ever be identified in retrospect. The retrieval process begins with a cue, but the rememberer is not aware of this process until the memory is activated. Moreover, there is no sense in asking whether the retrieved memory is the one targeted by the cue; there is no opportunity for it to be otherwise.

There is also no sense to be made of the idea of a cue failing to activate a memory during involuntary retrieval. If nothing is activated, then there is no cue – and also no memory. This is depicted in the leftmost column of the retrieval flowchart B&M provide in Figure 1. It is, however, strange to depict it at all. It is somewhat misleading to characterize proceeding straight down that column – answering *no* to the intentionality of retrieval, access to content, and a feeling of familiarity – as resulting in “no memory.” It is true that there is no memory in such cases, but that seems to be because there is no retrieval process at all. Nothing happens.

If this is the case, then it raises further problems for how we understand the depiction of involuntary memory and déjà vu in B&M’s model. In terms of what is consciously accessible, the starting point in cases of involuntary memory and déjà vu is not the cue, but the activation of a memory. The modification required for involuntary memory may be minimal, as these cases involve consciously accessing content. How B&M will accommodate this for déjà vu is less clear.

The complicated relationship between involuntary memories and their cues is further illustrated by cases where voluntary and involuntary retrieval are intermixed. B&M combine voluntary and involuntary memory into a single model, but they keep each form of retrieval distinct. Phenomena such as *chaining* (e.g., Mace, 2006; Pilotti, Chodorow, & Tan, 2004) suggest that a complete model should also account for interactions between these forms of retrieval. Chaining begins with voluntary memory, when the attempt to retrieve the targeted memory activates other memories inadvertently. An instance of voluntary remembering is thus chained to one or more instances of involuntary remembering.

Such a case nicely illustrates the distinct role cues play in voluntary and involuntary memory. In a case of chained remembering, the voluntary memory and the involuntary memory share a cue. Their relation to this cue, however, is different. The voluntary memory is targeted by that cue; the involuntary memory is not. The cue is understood *as a cue* for the voluntary memory from the outset. It only becomes a cue for the involuntary memory once that memory is activated. If the involuntary memory had not been activated, then it would not be a cue for that memory. Not so for the voluntary memory: Even if the retrieval attempt failed, the cue remains a cue.

B&M are right to push for greater attention to involuntary memory processes and a comprehensive account of retrieval. As this brief exploration of cues demonstrates, however, we think there is further work yet to be done.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Mace, J. H. (2006). Episodic remembering creates access to involuntary conscious memory: Demonstrating involuntary recall on a voluntary recall task. *Memory, 14*(8), 917–924.

Pilotti, M., Chodorow, M., & Tan, J. (2004). Asymmetric interference in implicit memory: Effects of study-test awareness and stimulus order. *The American Journal of Psychology, 117*(1), 43–64.

The relation of subjective experience to cognitive processing

Bennett L. Schwartz^a and Ali Pournaghдали^b

^aDepartment of Psychology, DM 283, Florida International University Miami, Miami, FL, USA and ^bLeonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

schwartzb@fiu.edu

pournagh@usc.edu

<https://www.bennettschwartz.com>

doi:10.1017/S0140525X23000262, e375

Abstract

Barzykowski and Moulin argue that common memory processes form the basis of involuntary autobiographical memory and the déjà vu experience. We think that they underemphasize the potential dissociability between processes that enact retrieval and the processes that produce conscious experience. We propose that retrieval and conscious experience result from different processes in both involuntary autobiographical memory and déjà vu experiences.

Barzykowski and Moulin (B&M) postulate that there is a relation between the déjà vu experience and involuntary autobiographical memory. We found their article to be rich in important insights into the nature of human memory as well as a valuable review of two important phenomena. However, we focus our criticism on the logic behind the rationale for equating the two phenomena. To quote from their abstract, B&M write “we hypothesize that both can be described as ‘involuntary’ or spontaneous cognitions, where IAMs deliver content and feeling of retrieval and déjà vu delivers only the feeling of retrieval.” Although we understand that the bulk of their paper explores the common retrieval processes that occur during involuntary autobiographical memory and déjà vu, we argue that in order to understand both of these phenomena, one must separate the processes involved in memory retrieval and the processes involved in conscious experience. Referring to Tulving’s challenge to the Doctrine of Concordance (Schwartz, 1999; Tulving, 1989), we assert that too often we assume that conscious experience derives from the same processes that produce retrieval, but, in reality, the conscious experience is often the product of different processes, including metacognitive processes.

Tulving (1989) argued that research in cognitive psychology assumed that cognitive processes, behavior, and conscious experience are closely correlated. In reviewing past work, he argued that cognitive psychology opened up the field to study hidden underlying processes, but that the next step was to look at the nature of consciousness and how it relates to cognitive processes. He argued that research needed to explore the nature of conscious experience and where it did not line up with cognitive processing. For an example, he cites implicit memory in which retrieval occurs without conscious access. Building on Tulving’s (1989) framework,

Schwartz (1999) argued that, at least in some cases, that one set of cognitive processes may be responsible for retrieval, but a separate cognitive process may be responsible for the conscious experience that accompanies retrieval and output of retrieved memories.

In our work on tip-of-the-tongue states, we showed that processes that lead to failed retrieval are dissociable from the processes that produce tip-of-the-tongue states (Schwartz & Pournaghdali, 2021). Early seminal work on tip-of-the-tongue states (e.g., Brown & McNeill, 1966) worked on the assumption that the cause of the experience of the tip-of-the-tongue state was word-retrieval failure. However, our work, inspired by Tulving's (1989) approach, shows that the processes that produce the tip-of-the-tongue state are more related to the processing of the cue or question than to access to the target or answer. For example, the familiarity of the cue (i.e., the question or stimulus) influences the reported number of tip-of-the-tongue states. Metcalfe, Schwartz, and Joaquim (1993) showed that repeating the cue, but not repeating the target, led to more tip-of-the-tongue states. More recently, Lee, Pournaghdali, and Schwartz (2022) showed that more fluently processed faces led to more tip-of-the-tongue states for the target name. This shows that with tip-of-the-tongue states, the fluent processing of a cue leads to a greater chance of a subjective experience about the target. The example of tip-of-the-tongue states supports the need to examine the processes of conscious experience as not being identical to the process of retrieval, consistent with Tulving's (1989) challenge to the Doctrine of Concordance.

We suspect that with both déjà vu and involuntary autobiographical memory, the subjective experience may be dissociable from the memory processes. Indeed, with respect to autobiographical memory, Neisser et al. (2023) distinguished between retrieval from autobiographical memory and the experience of mental time travel. Neisser et al. point to the observation that some amnesic patients can retrieve events from their past, but they lack the subjective experience of mental time travel, which they call the "feeling of pastness." Neisser et al. argue that these feelings of pastness are better thought of as a metacognitive experience about their retrieved memories, and, thus, support a challenge to the Doctrine of Concordance, because they are separable from retrieval.

We argue that evaluating the nature of the dissociation between a conscious experience and memory retrieval, as well as the possible interaction between the two is of critical importance. In our work, we have used a multidimensional extension of signal detection theory called general recognition theory (Ashby & Soto, 2015). We showed dissociation between the conscious experience of perceiving a face and facial-expression recognition (Pournaghdali, Schwartz, Hays, & Soto, 2023). However, our results indicate that the two phenomena are associated in a way that, in the absence of conscious experience, the strength of face recognition processes weakens but remains significantly above chance level, suggesting very complex relations between the object- and meta-level processes. Following from our work, general recognition theory can be extended to study dissociations and interactions between conscious experience and memory retrieval. Based on this, we argue that this approach is a suitable framework to evaluate different aspects of the Doctrine of Concordance. More specifically, the nature of the dissociations between cognitive processes and conscious experience can be tested using general recognition theory with high precision.

Although B&M declare themselves "agnostic" on the relation on how experience maps onto process, we argue that exploring this relation is important. Returning to our wheelhouse, thinking

about the tip-of-the-tongue experience in terms of a model in which the retrieval failure and the phenomenology were not identical has led to great progress in understanding tip-of-the-tongue states (e.g., Huebert, McNeely-White, & Cleary, 2023). Thus, we think it is important to clarify the role of retrieval from autobiographical memory and the feeling of pastness in involuntary autobiographical memory and to clarify the role of familiarity processes in memory and the subjective experience of a déjà vu.

Financial support. This paper was not supported by any grants.

Competing interest. None.

References

- Ashby, F. G., & Soto, F. A. (2015). Multidimensional signal detection theory. In *The Oxford handbook of computational and mathematical psychology* (pp. 13–34). Oxford University Press.
- Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325–337. [https://doi.org/10.1016/S0022-5371\(66\)80040-3](https://doi.org/10.1016/S0022-5371(66)80040-3)
- Huebert, A. M., McNeely-White, K. L., & Cleary, A. M. (2023). On the relationship between tip-of-the-tongue states and partial recollective experience: Illusory partial recollective access during tip-of-the-tongue states. *Journal of Experimental Psychology: General*, 152(2), 542–570. <https://doi.org/10.1037/xge0001292>
- Lee, H., Pournaghdali, A., & Schwartz, B. L. (2022). Does clarity of famous faces affect the rate of tip-of-the-tongue states? Poster presented at Psychonomics Conference, Boston, MA, November 19, 2022.
- Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 851–861. <https://doi.org/10.1037/0278-7393.19.4.851>
- Neisser, J., Abreu, G., Drane, D.L., Pederson, N.P., Parsons, T. D., Cleary, A. M., & McNeely-White, K. L. (2023). Opening a conceptual space for metamemory experience. *New Ideas in Psychology*, 69, 100995.
- Pournaghdali, A., Schwartz, B. L., Hays, J., & Soto, F. A. (2023). Sensitivity vs. awareness curve: A novel model-based analysis to uncover the processes underlying nonconscious perception. *Psychonomic Bulletin & Review*, 30, 553–536. <https://doi.org/10.3758/s13423-022-02180-3>
- Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin & Review*, 6, 379–393. <https://doi.org/10.3758/BF03210827>
- Schwartz, B. L., & Pournaghdali, A. (2021). Tip-of-the-tongue-states: Past and future. In A. Cleary & B. L. Schwartz (Eds). *Memory quirks: The study of odd phenomena in memory*. Routledge Press. pp. 207–223.
- Tulving, E. (1989). Memory: Performance, knowledge, and experience. *European Journal of Cognitive Psychology*, 1, 3–26.

The misidentification syndromes and source memory deficits with their neuroanatomical correlates from neuropsychological perspective

Rafał Sikorski^{a,b}  and Emilia J. Sitek^{a,c} 

^aLaboratory of Clinical Neuropsychology, Neurolinguistics and Neuropsychotherapy, Department of Neurological and Psychiatric Nursing, Faculty of Health Sciences, Medical University of Gdansk, ul. Debinki 7, Gdansk, Poland; ^bDepartment of Neurological Rehabilitation, St. Vincent Hospital, Pomeranian Hospitals, Ul. Wójta Radtkego 1, Gdynia, Poland and ^cDepartment of Neurology, St. Adalbert Hospital, Copernicus PL, Al. Jana Pawła II 50, Gdańsk, Poland

rafal.sikorski@gumed.edu.pl

emilia.sitek@gumed.edu.pl

<https://structure.mug.edu.pl/520>

<https://structure.mug.edu.pl/520>

doi:10.1017/S0140525X23000274, e376

Abstract

The suggested model is discussed with reference to two clinical populations with memory disorders – patients with misidentification syndromes and those with source memory impairment, both of whom may present with (broadly conceived) déjà vu phenomenon, without insight into false feeling of familiarity. The role of the anterior thalamic nucleus and retrosplenial cortex for autobiographical memory and familiarity is highlighted.

Barzykowski and Moulin (B&M) presented an interesting model of autobiographical memory integrating both involuntary autobiographical memories (IAMs) and déjà vu phenomena. We suggest analysing déjà vu phenomenon in a broader context of misidentification syndromes (MIS) and propose the inclusion of anterior thalamic nuclei (ANT), particularly relevant for source memory, and retrosplenial cortex (RSC) into the neuroanatomical basis of the model.

Firstly, the model proposed by B&M suggests the sequential nature of memory retrieval process as depicted in [Figure 1](#). Drawing upon a clinical example of delusional MIS, we suggest that these processes are not necessarily sequential. While in Capgras syndrome (CS) access to memory content is preserved (patient's proxy face is correctly recognized), feeling of familiarity is missing (the proxy is perceived as an imposter). In contrast, patients with Fregoli syndrome misrecognize strangers as proxies, presenting with false feeling of familiarity. A recent case of highly selective CS supported the idea that right temporal lesions lead to dissociation of a familiar face with its emotional value (Nuara et al., 2020), which is in line with double dissociation between overt face recognition (fusiform face area) and autonomic recognition (limbic cortex) hypothesized by Ellis and Lewis (2001). Of note, Darby, Laganier, Pascual-Leone, Prasad, and Fox (2017) recently linked familiarity to left RSC and the belief evaluation to the right frontal cortex. The nearest posterior cingulate cortex was already linked to specific familiarity processing (Qin et al., 2012) and autobiographical memory (Spreng, Mar, & Kim, 2009). Thus, as RSC is active during the retrieval of all types of autobiographical information – both emotionally laden and neutral (Vann, Aggleton, & Maguire, 2009) and linked to the limbic system (Powell et al., 2018) we suggest its incorporation to the neuroanatomical basis of B&M's model.

As Capgras and Fregoli syndromes (or their analogues for place memory, e.g., reduplicative paramnesia) may be regarded as counterparts of jamais vu and déjà vu phenomena, it is surprising that the proposed model does not address the emotional component of familiarity with reference to relevant neural networks and that the other side of the coin – jamais vu phenomenon – is not thoroughly discussed.

We strongly believe that feeling of familiarity has its emotional aspect and cannot be reduced to (meta)-cognitive components. A broad meta-analysis of functional magnetic resonance imaging or positron emission tomography data involving a familiarity task by Horn et al. (2016) evidenced the dependence of the involvement of emotional networks on the type of paradigm used. The first two, fully laboratory-based, paradigms are based on the initial encoding phase of stimuli that are not specifically familiar to the participant, and they activate only the prefrontal and parietal regions. However, the third paradigm, based on stimuli previously known to the participant is assumed to refer to self-related emotions and personal experience, activates the aforementioned regions and also parts of the limbic system (Horn et al., 2016).

That revealed the undoubted role of emotional processing of stimuli of specific familiarity, and the paradigm used seems closer to the complexity of everyday functioning and potentially directly covers a larger group of clinical cases, such as the aforementioned delusional MIS.

The reduction of autobiographical memory model to its (meta)-cognitive components and the lack of inclusion of emotional networks may affect its explanatory value in the clinic, for example, in the case of MIS. Not only does a conviction that a known person or place has been replaced by its copy have emotional consequences for the affected person, but also this conviction itself is based on the impairment of emotional processing. It seems to us that déjà vu is also likely to be emotionally loaded. Thus, we postulate that the autobiographical memory model would benefit from emphasizing the role of neural networks that integrate memory for facts with their emotional value. The interactions between medial temporal-lobe structures and prefrontal medial cortex could have been given more importance.

Furthermore, the suggested model focusing on the normative aspects of déjà vu fails to include a stage of belief evaluation, leading to true versus false feeling of familiarity. Implementing a module reflecting belief evaluation into it, as operationalized by Darby et al. (2017), could improve the model's explanatory value in the memory clinic, especially in the context of MIS.

Secondly, we suggest further refinement of the neuroanatomical basis of déjà vu. In [Table 1](#), B&M referred to medial temporal-lobe circuitry and cingulate cortex/prefrontal cortex. We would like to pinpoint the important role of ANT for source memory (particularly relevant for déjà vu phenomenon), as demonstrated in research on Korsakoff syndrome. The most characteristic symptom of Korsakoff's is severe declarative memory impairment with the prominent loss of the order of events over time (achronogenesis) or its temporal context (Kopelman, 2015). Temporal encoding deficit in Korsakoff syndrome can occur independently of typical frontal pathology (Dillingham, Milczarek, Perry, & Vann, 2021). Patients reveal distinct difficulties with source memory – the less efficient encoding of contextual temporal and also spatial information. While mammillary bodies specialize in direct and integrated action with ANT on recollective-based recognition, other diencephalic nuclei do not show an exclusive specialization in recognition. Nevertheless, ANT are more closely related to recall memory and the medial dorsal thalamic nuclei are hypothesized to be related to recognition memory (Aggleton, Dumont, & Warburton, 2011). The crucial role of the ANT in amnesia linked to Korsakoff's syndrome, albeit hypothesized for years, was clearly evidenced over two decades ago (Harding et al., 2000). Recently, Segobin et al.'s (2019) investigation using diffusion tensor imaging sequence analysis revealed that the disconnection between ANT and hippocampus, leading to ANT atrophy, was a neuroimaging marker of thalamic amnesia (Segobin et al., 2019).

Finally, we propose integrating the model by Barzykowski and Moulin with the recent Aggleton and O'Mara (2022), which assumes a parallel and partially convergent operation of two memory streams: Hippocampal–cortical and a medial diencephalic–cortical stream. Hippocampus remains the core element of the first stream, while ANT are at the heart of the second one. The interactions between these streams seem crucial for mnemonic consolidation (Aggleton & O'Mara, 2022).

Financial support. None.

Competing interest. None.

References

- Aggleton, J. P., Dumont, J. R., & Warburton, E. C. (2011). Unraveling the contributions of the diencephalon to recognition memory: A review. *Learning & Memory, 18*(6), 384–400. <https://doi.org/10.1101/lm.188461>
- Aggleton, J. P., & O'Mara, S. M. (2022). The anterior thalamic nuclei: Core components of a tripartite episodic memory system. *Nature Reviews Neuroscience, 23*(8), 505–516. <https://doi.org/10.1038/s41583-022-00591-8>
- Darby, R. R., Laganieri, S., Pascual-Leone, A., Prasad, S., & Fox, M. D. (2017). Finding the imposter: Brain connectivity of lesions causing delusional misidentifications. *Brain, 140*(2), 497–507. <https://doi.org/10.1093/brain/aww288>
- Dillingham, C. M., Milczarek, M. M., Perry, J. C., & Vann, S. D. (2021). Time to put the mammillothalamic pathway into context. *Neuroscience & Biobehavioral Reviews, 121*, 60–74. <https://doi.org/10.1016/j.neubiorev.2020.11.031>
- Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: A window on face recognition. *Trends in Cognitive Sciences, 5*(4), 149–156. [https://doi.org/10.1016/s1364-6613\(00\)01620-x](https://doi.org/10.1016/s1364-6613(00)01620-x)
- Harding, A., Halliday, G., Caine, D., & Kril, J. (2000). Degeneration of anterior thalamic nuclei differentiates alcoholics with amnesia. *Brain: A Journal of Neurology, 123*(1), 141–154. <https://doi.org/10.1093/brain/123.1.141>
- Horn, M., Jardri, R., D'Hondt, F., Vaiva, G., Thomas, P., & Pins, D. (2016). The multiple neural networks of familiarity: A meta-analysis of functional imaging studies. *Cognitive, Affective and Behavioral Neuroscience, 16*(1), 176–190. <https://doi.org/10.3758/s13415-015-0392-1>
- Kopelman, M. D. (2015). What does a comparison of the alcoholic Korsakoff syndrome and thalamic infarction tell us about thalamic amnesia? *Neuroscience & Biobehavioral Reviews, 54*, 46–56. <https://doi.org/10.1016/j.neubiorev.2014.08.014>
- Nuara, A., Nicolini, Y., D'Orto, P., Cardinale, F., Rizzolatti, G., Avanzini, P., ... De Marco, D. (2020). Catching the imposter in the brain: The case of Capgras delusion. *Cortex, 131*, 295–304. <https://doi.org/10.1016/j.cortex.2020.04.025>
- Powell, A. L., Hindley, E., Nelson, A. J., Davies, M., Amin, E., Aggleton, J. P., & Vann, S. D. (2018). Lesions of retrosplenial cortex spare immediate-early gene activity in related limbic regions in the rat. *Brain and Neuroscience Advances, 2*, 239821281881123. <https://doi.org/10.1177/2398212818811235>
- Qin, P., Liu, Y., Shi, J., Wang, Y., Duncan, N., Gong, Q., ... Northoff, G. (2012). Dissociation between anterior and posterior cortical regions during self-specificity and familiarity: A combined fMRI-meta-analytic study. *Human Brain Mapping, 33* (1), 154–164. <https://doi.org/10.1002/hbm.21201>
- Segobin, S., Laniece, A., Ritz, L., Lannuzel, C., Boudehent, C., Cabé, N., ... Pitel, A. L. (2019). Dissociating thalamic alterations in alcohol use disorder defines specificity of Korsakoff's syndrome. *Brain, 142*(5), 1458–1470. <https://doi.org/10.1093/brain/awz056>
- Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience, 21*(3), 489–510. <https://doi.org/10.1162/jocn.2008.21029>
- Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience, 10*(11), 792–802. <https://doi.org/10.1038/nrn2733>

What do we gain (or lose) by considering déjà vu a part of autobiographical memory?

Søren Risløv Staugaard 

Arts Studies, Aarhus University, Aarhus C, Denmark
sorsta@au.dk

doi:10.1017/S0140525X23000043, e377

Abstract

I argue the relationship between déjà vu and autobiographical memory is not continuous, but more akin to a path diagram. The starting points might be overlapping, but eventually the paths diverge dependent on whether there is memory content to be retrieved. I am worried that considering déjà vu as part of autobiographical memory solves more problems than it creates.

Barzykowski and Moulin present a compelling case for why déjà vu could be considered a special case of involuntary

autobiographical memory, more specifically one in which retrieval was initiated but never completed. While I find the arguments and evidence convincing, I cannot let go of the “why.” Why take a conceptually and phenomenologically discrete concept such as déjà vu and redefine it as failed retrieval? I believe the authors run the risk of introducing problems that did not exist before, and it is not immediately obvious to me what the gains are.

Considering déjà vu and autobiographical memory on a continuum assumes that one can become the other with some additional components or higher intensity of activation. In line with this, the authors argue that additional activation of traces (bottom-up) or additional elaboration of details (top-down) can transform déjà vu into autobiographical memory proper. This premise begs the very intriguing question of when a memory begins and when it is completed, and what types of phenomena should be included at each end of this theoretical line.

The sensation of déjà vu may at first feel close enough to a memory that the individual initiates a deliberate search, only to decide there was no memory after all – it was a trick of the mind. An auditory hallucination could be considered analogous: An individual hears their own name and searches their apartment for the source, only to discover that no one is around. However, no amount of additional search or neurophysiological activation will lead to a source, because there is none. I believe this is also true for déjà vu at a conceptual and phenomenological level. It might also be true at a neurophysiological level, but that is an empirical question that might never be solved.

An autobiographical memory is the conscious recollection of a past event. This is the generally agreed upon definition referred to by the authors on page 6. Activation, or sensation, or any process that does not lead to recollection of a past event is not an autobiographical memory by this definition. Electronic stimulation of the temporal lobe is not autobiographical memory, and a focal temporal-lobe epileptic seizure is not an autobiographical memory. Why then award this special status to déjà vu? On the one hand, the authors base their argument on the shared neurophysiological correlates of both déjà vu and involuntary autobiographical memory, but on the other hand, they explicitly state that déjà vu and autobiographical memory branch away from each other as a result of a conscious evaluation of plausibility (target article, sect. 3, para. 1). On the one hand, déjà vu can be considered part of autobiographical memory because of the shared neurophysiological correlates, but at the same time, the conscious evaluation of plausibility is essential to the phenomenon – and the evaluation is that déjà vu is emphatically not a memory. In this way, déjà vu comes to present a problem for the (conventional) definition of autobiographical memory, if we want to include it under this definition.

Phenomenologically speaking, déjà vu is a satisfactory explanation for a distinct experience, analogous to auditory hallucination being a satisfactory explanation for hearing your cell phone “ding,” when there is in fact no message. When the individual says, “oh, it wasn't a memory, it was just déjà vu,” they are correct in an ontological sense. To say they actually experienced failed retrieval of an autobiographical memory appears to introduce an unnecessary extra step to this experience and to be at odds with the phenomenology of the experience.

Déjà vu, similarly to other tricks of the mind, could arise from a Bayesian principle in consciousness (Ramachandran, 1998): If the current situation shares a critical amount of features with a stored configuration of features, the most likely explanation is that the situation is familiar rather than occurring completely

by chance. However, the stored configuration of features could refer to one or more specific past events, some fragments of multiple events, repeated events, or potentially nothing relevant at all. Importantly, there is no way of knowing (the hard problem of consciousness) and besides, the whole thing is satisfactorily resolved as the error it most likely is: A déjà vu.

I would argue that the relationship between déjà vu and autobiographical memory is not continuous, but more akin to a path diagram, where the starting points might be overlapping (pattern matching in the hippocampus), the next step might be overlapping in some situations (a sense of familiarity), but when this activation reaches consciousness, the paths diverge. At this point, one path leads to not-memory (déjà vu), while another path leads to memory (autobiographical memory). Redefining déjà vu as failed retrieval conjures up the ghost of a hypothetical subconscious memory, which introduces the hard problem of consciousness. What problem does it solve?


Financial support. This commentary received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

Reference

Ramachandran, V. S. (1998). Consciousness and body image: Lessons from phantom limbs, Capgras syndrome and pain asymbolia. *Philosophical Transactions of the Royal Society B*, 353, 1851–1859.

Déjà vu: A botched memory operation, illegitimate to start with

Debora Stendardi^a, Anindita Basu^b, Alessandro Treves^b and Elisa Ciaramelli^a 

^aDipartimento di Psicologia ‘Renzo Canestrari’, Università di Bologna, Bologna, Italy and ^bSISSA, Cognitive Neuroscience, Trieste, Italy
debora.stendardi2@unibo.it
abasu@sissa.it
ale@sissa.it
elisa.ciaramelli@unibo.it

doi:10.1017/S0140525X2300016X, e378

Abstract

Rather than a natural product, a computational analysis leads us to characterize déjà vu as a failure of memory retrieval, linked to the activation in neocortex of familiar items from a compositional memory in the absence of hippocampal input, and to a misappropriation by the self of what is of others.

Freud (1901) had already noted that déjà vu involves memory retrieval: “the uncanny feeling we have, in certain situations, of having had exactly the same experience once before, or of having once before been exactly in the same place, though our efforts never succeed in remembering the previous occasion that announces itself this way.”

Barzykowski and Moulin’s (B&M’s) analytical effort conceptualizes déjà vu as one possible output of a continuum (although

structured as a discrete decision tree in Fig. 1 of their article) of spontaneous memory retrieval phenomena: While involuntary memories would be the unexpected retrieval of content, déjà vu would be that of an unjustified feeling of familiarity (not accompanied by a particular memory in mind). Déjà vu, indeed, soon reveals itself as the phantom of a memory, one we do not belong to, a “feeling of retrieval” to say it with the authors, but one that people describe as disquieting, eerie, awkward (Brown, 2003).

What exactly happens during déjà vu that disorients us? How can a spontaneous retrieval process go awry? According to the authors, déjà vu happens when there is a feeling of familiarity that does not pass a plausibility check (Fig. 1 of their article), and it is on the neural mechanism of this feeling of familiarity and the question of its *implausibility* (*memory absence*) that our commentary is focused.

Many times we experience familiarity of unclear origin without having a déjà vu. The butcher-on-the-bus phenomenon, where someone feels familiar but it is not clear from where, is one such case, but it is not disquieting an experience. On the other hand, sometimes déjà vu occurs in familiar places or involves familiar people, where feelings of familiarity would be plausible and justified. Moreover, it is not clear how the dichotomous (explicit?) plausibility signal in Figure 1 can be computed in the absence of access to content.

We think that the feeling of familiarity that accompanies the experience of déjà vu (troubles us because it) is fundamentally different from other instances of familiarity mentioned in the paper: It is not relative to a single item, but to a composition of items, to an experience, albeit fragmented: The place, who was there, some words we uttered, something that will happen next. Hence, we expect the ensuing recollection of the corresponding event that instead does not happen. Perhaps that feels implausible: To have forgotten an entire event that we are currently reliving.

Our recent modeling study (Ryom, Stendardi, Ciaramelli, & Treves, 2022) offers a computational explanation of associative retrieval failures. These are in fact very frequent, especially if retrieval is triggered by the activation of partial cues in the neocortex, rather than by hippocampal activity indexing memory. Our model network is comprised of “Potts units,” which represent patches of cortex, interacting through long-range connections (Fig. 1). A compositional memory, such as the memory for a complex event (e.g., my dog hid my friend’s sweater in the park), is conceived as composed of several items, each of which has a pre-established neocortical representation (dog, park, sweater). Storing this new memory only involves acquiring the novel connections among participating items. Memory retrieval could be triggered either by the activation of a partial cue in the cortex, which is a variable fraction of the units active in the memory (e.g., sweater + friend), or by a hippocampal input that sustainedly cues all the memory units simultaneously, working as an index to the compressed representation of the entire memory (see Fig. 1). One main finding of our study is that the cortical storage capacity for compositional memories is much lower than previously calculated for unitary representations (Treves & Rolls, 1994). The reason is that while the hippocampus is thought to store newly assembled compressed representations of each episode in memory, the neocortex has to make do with reusing pre-established representations of the various components of the episode (Ciaramelli, Lauro-Grotto, & Treves, 2006). The ability of the neocortical network to retrieve compositional memories from partial cues, in the absence of hippocampal input, is shown, analytically and with computer simulations, to be severely limited, plagued by the interference from competing representations (Ryom et al., 2022).



Figure 1 (Stendardi et al.). The hippocampus activates all the five items constituting the real event “my dog hid my friend’s sweater in the park” (straight gray arrows). The activation of two highly familiar items in the absence of hippocampal input may result in déjà vu (fragmented red arrows). Each item has a sparse distributed but partially localized representation over the cortex.

On this view, déjà vu could be characterized as an “incomplete” memory state where some familiar items from a compositional memory (or from several distinct memories) get activated in neocortex (e.g., kids + bench; Fig. 1), in the absence of hippocampal input. This activation is sufficient to trigger familiarity for an experience, but not the reinstatement of a full-fledged memory (assuming one exists). The ensuing feeling of familiarity may be particularly uncanny if the partial cue activates self-relevant items or schemata in the neocortex (Stendardi, Biscotto, Bertossi, & Ciaramelli, 2021), conferring self-relevance to a memory that might potentially be false, and should last until activated memory fragments are enough to finally trigger monitoring mechanisms that explicitly refute the participation of the self; or, we suggest, the (false) memory can be abandoned based on an implicit network signal that automatically reads out high levels of simultaneous activity in the neocortex that in the absence of hippocampal activity are more compatible with imagination than with memory. By contrast, the protracted failure of memory monitoring may lead to confabulation, the false memory for unhappened events (Gilboa et al., 2006; see also Moulin, 2013). Similar to déjà vu, confabulation entails fragments of memory traces, and is mostly self-related (Gilboa et al., 2006; see also Moulin, 2013). Unlike déjà vu, confabulation is not abandoned but endorsed confidently. Interestingly, confabulation is triggered by familiar stimuli (Ciaramelli, 2008), and dampened by reducing the cognitive resources available for assembling (wrongly) memory elements (Ciaramelli, Ghetti, & Borsotti, 2009).

Does the activation of multiple (self-relevant) memory fragments make déjà vu so unique and distinguishable from other illusory familiarity phenomena? Is the estranging feeling

associated with déjà vu the by-product of a just foiled risk of confabulation? Future studies should test this hypothesis, for example, studying whether déjà vu is associated with the activation of ventral prefrontal cortex regions, and the computational conditions conducive to memory or confabulatory signals.



Financial support. The authors acknowledge the support of a PRIN grant from the Italian Ministry of Education, University, and Research to E. C. and A. T. (PRIN #20174TPEFJ).

Competing interest. None.

References

- Brown, A. S. (2003). A review of the déjà vu experience. *Psychological Bulletin*, 129, 394–413.
- Ciaramelli, E. (2008). The role of ventromedial prefrontal cortex in navigation: A case of impaired wayfinding and rehabilitation. *Neuropsychologia*, 46, 2099–2105.
- Ciaramelli, E., Ghetti, S., & Borsotti, M. (2009). Divided attention during retrieval suppresses false recognition in confabulation. *Cortex*, 45, 141–153.
- Ciaramelli, E., Lauro-Grotto, R., & Treves, A. (2006). Dissociating episodic from semantic access mode by mutual information measures: Evidence from aging and Alzheimer’s disease. *Journal of Physiology-Paris*, 100, 142–153.
- Freud, S. (1901). *The psychopathology of everyday life* (Standard Edition, Vol. VI). Hogarth Press.
- Gilboa, A., Alain, C., Stuss, D. T., Melo, B., Miller, S., & Moscovitch, M. (2006). Mechanisms of spontaneous confabulations: A strategic retrieval account. *Brain*, 129, 1399–1414.
- Moulin, C. J. (2013). Disordered recognition memory: Recollective confabulation. *Cortex*, 49, 1541–1552.
- Ryom, K. I., Stendardi, D., Ciaramelli, E., & Treves, A. (2022). Computational constraints on the associative recall of spatial scenes. *bioRxiv*.
- Stendardi, D., Biscotto, F., Bertossi, E., & Ciaramelli, E. (2021). Present and future self in memory: The role of vmPFC in the self-reference effect. *Social Cognitive and Affective Neuroscience*, 16, 1205–1213.
- Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4, 374–391.

Involuntary autobiographical memories and déjà vu: When and why attention makes a difference

Manila Vannucci^a  and Maciej Hanczakowski^b 

^aDepartment of Neuroscience, Psychology, Drug Research and Child Health (NEUROFARBA) – Section of Psychology, University of Florence, Firenze, Italy and ^bFaculty of Psychology and Cognitive Science, Adam Mickiewicz University, Poznań, Poland

manila.vannucci@unifi.it

maciej.hanczakowski@gmail.com

<https://www.unifi.it/p-doc2-2022-0-A-2b333d2b3928-0.html>

lamelab.org

doi:10.1017/S0140525X23000031, e379

Abstract

The target article claims that involuntary autobiographical memories and déjà vu are based on the same retrieval processes, although they result in different phenomenological states. Here we argue that the differential engagement of attention at various stages of memory may be one of the determinants of when common retrieval processes give rise to such different experiences.

The main claim of the target article is that involuntary autobiographical memories (IAMs) and déjà vu are two facets of spontaneous memory, based on the same retrieval processes. The authors argue that empirical and theoretical advances can be spurred by grouping these two phenomena under one conceptual framework of spontaneous memory-related cognition. While we agree that IAMs and déjà vu build on common memory processes, we argue that considering them together can move the area forward by highlighting their differences rather than similarities. The most obvious difference is that while IAMs are experienced on a daily basis, déjà vu is a rather esoteric experience eluding people most of the time. We thus call for a line of research identifying how the common memory processes underpinning IAMs and déjà vu can result in vastly differing phenomenological manifestations, and postulate that non-memory mechanisms must be considered to understand this.

One such mechanism, which the target article mentions only briefly, is attention. Attention operates throughout the memory processes, from cue encoding, through retrieval, up to attributional processes. Although the interplay of attention and memory has long been of interest (Craig, Govoni, Naveh-Benjamin, & Anderson, 1996), the investigation of attention in relation to both IAMs and déjà vu is much more recent, reflecting also a relatively newer focus on spontaneous memory processes more generally. Out of these two phenomena, more empirical work has been devoted to outlining the role of attention in IAMs. This has been investigated both in naturalistic settings and in laboratory, with both lines of research demonstrating how IAMs are more likely to occur during activities that require little focused attention or concentration (e.g. Berntsen & Hall, 2004; Kvavilashvili & Mandler, 2004; Schlagman & Kvavilashvili, 2008).

While the importance of diffuse attention for experiencing IAMs is without doubt, the exact mechanisms by which attention impacts spontaneous memory are still debated. Vannucci, Hanczakowski, Pelagatti, and Chiorri (2019) provided evidence that reduced

incidence of IAMs with an increase of the attentional load is due to both IAMs being less likely to be retrieved and less likely to be noticed – or cross what Barzykowski and Moulin (B&M) refer to as the awareness threshold. Regarding the retrieval mechanism, recently Barzykowski, Hajdas, Radel, and Kvavilashvili (2022) suggested that high attentional load reduces the chances that potential environmental triggers of IAMs would be noticed, resulting not only in lowered incidence of retrieval of IAMs but also later in impaired recognition of these potential cues. Regarding the post-retrieval mechanism, the increased attentional load implemented by Vannucci et al. resulted in reduced ratings of meta-awareness of IAMs, which were collected when participants' task was interrupted to make participants focus instead on the contents of their minds in the moments just preceding the interruption.

Thus, the acknowledgment that attention plays a role in both eliciting and experiencing IAMs is key here, showing that IAMs are as much a result of memory processes as of attributions following retrieval. This complex picture is acknowledged by B&M, but their treatment of post-retrieval processes in relation to IAMs remains limited. The authors focus mostly on how “phenomenologically sound” memories are more likely to cross the awareness threshold while mostly neglecting how non-memory factors, such as attention, affect this threshold in the first place. The relatively sparse treatment of these issues is highlighted by the fact that the flowchart in Figure 1 includes a specific post-retrieval step for experiencing déjà vu (“Is this plausible?”) but not for IAMs, where having spontaneous access to memory contents directly leads to involuntary memories.

Interestingly, B&M do acknowledge the role of attention when discussing the awareness threshold in relation to déjà vu. They suggest that these states “may be more likely to arise when one is engaged in attention demanding activities,” which may result in an elevated awareness threshold and thus reduced chances of experiencing full recollection that would preclude déjà vu. This is an interesting suggestion as it directly pits déjà vu against IAMs, for which meta-awareness occurs precisely when one is engaged in less demanding activities (Vannucci et al., 2019). If borne out, this hypothesis would thus directly implicate attention as a “switch” mechanism responsible for whether common retrieval processes may result in such different states.

But is the involvement of attention in déjà vu limited to the appraisal and attribution of the products of retrieval? For IAMs, such involvement seems to start earlier, when the retrieval process is preceded by encoding of retrieval cues (Barzykowski et al., 2022). While research on cue processing in déjà vu states is less developed, some work does point to involvement of attentional processes also at this stage. As shown by Brown and Marsh (2009), a brief glance at an unfamiliar symbol increased feelings of knowing the stimulus from pre-experimental exposure when the stimulus was fully processed moments later. These and other similar demonstrations of a false sense of prior experience following a brief encounter with the stimulus (e.g. Jacoby & Whitehouse, 1989; Klinger, 2001) suggest that déjà vu may be more likely to occur when someone is engaged in attention-demanding activities and the processing of peripheral information – potential memory cues – is superficial. Again, this contrasts directly with the studies indicating that attention-demanding activities reduce the likelihood of IAMs being elicited. Thus, just as in the case of post-retrieval processing, attention may serve as a switch determining how potential retrieval cues are processed, which in turn determines whether these cues elicit retrieval as déjà vu versus IAM.

To summarize, differential engagement of attention at various stages of memory may be one of the determinants of whether common retrieval processes give rise to contextually rich episodic memories experienced as IAMs or spurious feelings of familiarity experienced as déjà vu. The full framework of these spontaneous manifestations of memory processing must involve an in-depth understanding of non-memory mechanisms, of which attention is likely to be just one example, to understand how a common process of matching cues to the contents of memory can result in such vastly different phenomenological states.


Financial support. This work received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

- Barzykowski, K., Hajdas, S., Radel, R., & Kvavilashvili, L. (2022). Effects of inhibitory control capacity and cognitive load on involuntary past and future thoughts: A laboratory study. *Consciousness and Cognition*, 102, 103353. doi: 10.1016/j.concog.2022.103353
- Berntsen, D., & Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Memory & Cognition*, 32, 789–803. doi: 10.3758/BF03195869
- Brown, A. S., & Marsh, E. J. (2009). Creating illusions of past encounter through brief exposure. *Psychological Science*, 20, 534–538. doi: 10.1111/j.1467-9280.2009.02333
- Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125, 159–180. doi: 10.1037/0096-3445.125.2.159
- Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, 118, 126–135. doi: 10.1037/0096-3445.118.2.126
- Klinger M. R. (2001). The roles of attention and awareness in the false recognition effect. *American Journal of Psychology*, 114, 93–114. doi: 10.2307/1423382
- Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology*, 48, 47–94. doi: 10.1016/S0010-0285(03)00115-4
- Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition*, 36, 920–932. doi: 10.3758/mc.36.5.920
- Vannucci, M., Hanczakowski, M., Pelagatti, C., & Chiorri, C. (2019). Visual attentional load affects the frequency of involuntary of autobiographical memories and their level of meta-awareness. *Memory and Cognition*, 47, 117–129. doi: 10.3758/s13421-018-0854-0

A spontaneous neural replay account for involuntary autobiographical memories and déjà vu experiences

Mohith M. Varma and Rongjun Yu 

Department of Management, School of Business, Hong Kong Baptist University, Hong Kong, S.A.R. China
rongjunyu@hkbu.edu.hk

doi:10.1017/S0140525X23000109, e380

Abstract

Barzykowski and Moulin argue both involuntary autobiographical memories and déjà vu experiences rely on the same involuntary memory retrieval processes but their underlying neurological basis remains unclear. We propose spontaneous neural replay in the default mode network (DMN) and hippocampus as the basis for involuntary autobiographical memories, whereas for déjà vu experiences such transient activation is limited to the DMN.

Involuntary autobiographical memories (IAMs) and déjà vu experiences are not uncommon in our daily life. Both involve memory recall that occur spontaneously without any deliberate/conscious intention to recall them. Barzykowski and Moulin (B&M) posit both IAMs and déjà vu experiences result from the same involuntary memory retrieval processes, while claiming that IAMs are distinct because of their explicit accessibility to the retrieved memory content. However, the neurological basis of these two phenomena remains unclear in the target article. In this commentary, we propose spontaneous neural replay in the hippocampus and default mode network (DMN) as the basis for experiencing IAMs, whereas for déjà vu experiences such transient activation is limited to the DMN.

Spontaneous neural replay refers to spontaneously generated reactivation of neural activity and configurations related to past experiences (Schuck & Niv, 2019). It occurs preferentially during brief bursts (~30–100 ms) of high-frequency (~200 Hz) neural firing known as sharp-wave ripples (Dickey et al., 2022). Such spontaneous brain activity occurs as a result of intrinsic brain dynamics as well as cued reactivation by sensory stimuli (Liu, Nour, Schuck, Behrens, & Dolan, 2022). Spontaneous neural replay is a critical process supporting memory consolidation (Deuker et al., 2013) and memory retrieval (Vaz, Wittig, Inati, & Zaghoul, 2020). Spontaneous neural replay has been noted to occur in the hippocampal (Gillespie et al., 2021) and the neocortical memory systems (e.g., DMN nodes; Higgins et al., 2021). Hippocampus and DMN nodes (e.g., medial prefrontal cortex [mPFC]; posterior cingulate cortex [PCC]) have been implicated in supporting a range of memory-related functions including autobiographical memory retrieval (Cabeza & St Jacques, 2007; Philippi, Tranel, Duff, & Rudrauf, 2015).

According to the *cascaded memory systems* model (Kaefer, Stella, McNaughton, & Battaglia, 2022), spontaneous neural replay in the DMN can be independently triggered in the absence of corresponding hippocampal replay. Transient activation limited to DMN and its propagation to areas downstream of DMN (e.g., unimodal sensory cortices) correspond to reactivation of only factorized feature representations that make up a memory (Kaefer et al., 2022). Such features may involve various attributes or details of a memory's content. Kaefer et al. (2022) also claim that hippocampal replay in conjunction with spontaneous neural replay in DMN and its downstream propagation to primary sensory cortices are essential to the retrieval of a complete memory. Being a critical brain region for memory encoding (Scoville & Milner, 1957), hippocampus is considered to generate and store highest-order memory representations (Kaefer et al., 2022). These highest-order memory representations could be viewed as index codes that support the binding of contextual details of a memory (Simons, Ritchey, & Fernyhough, 2022). Therefore, hippocampal replay reactivates the highest-order memory index codes necessary for binding or integrating the factorized feature representations of a memory into a cohesive whole (Goode, Tanaka, Sahay, & McHugh, 2020).

Inspired by the *cascaded memory systems* model (Kaefer et al., 2022), we propose spontaneous neural replay in the DMN occurring in the absence of hippocampal replay forms the neural basis of déjà vu (Fig. 1a). For déjà vu experiences, the memory retrieval process is incomplete and the memory contents are inaccessible because the memory representations spontaneously reactivated in the DMN are *factorized and lack the highest-order memory index codes* presumed to be generated and stored only in the hippocampus (Kaefer et al., 2022). Therefore, in the absence of

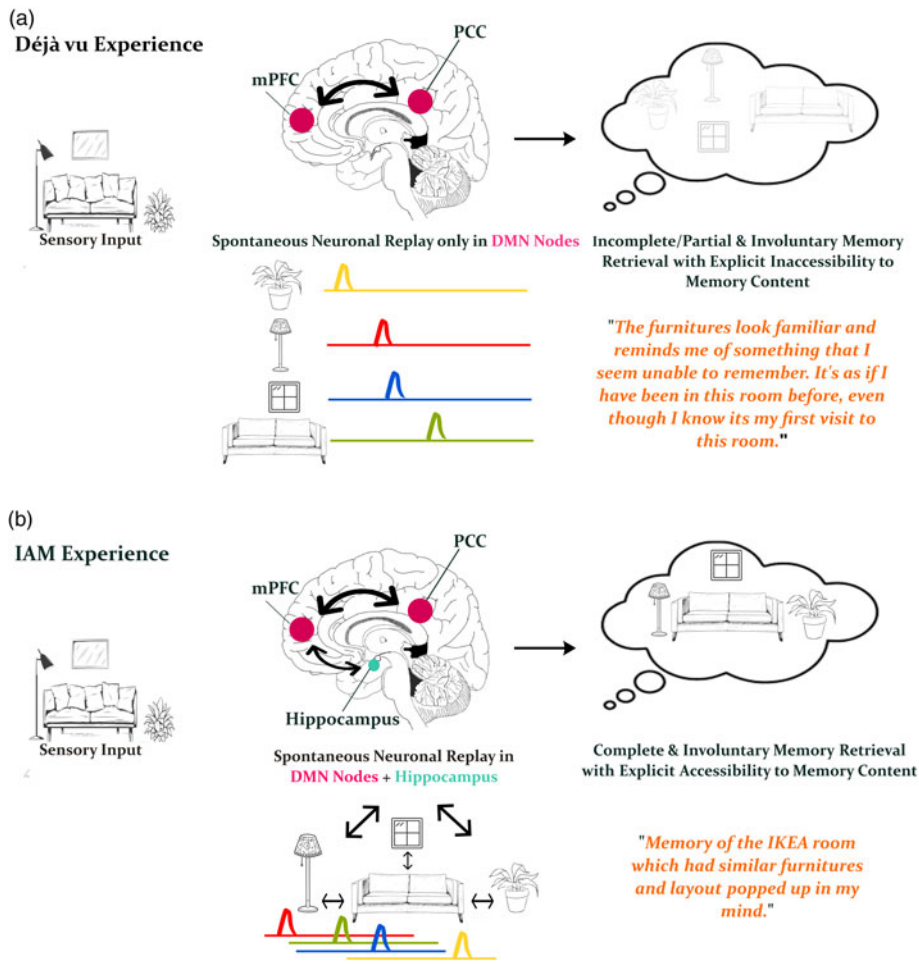


Figure 1 (Varma and Yu). The spontaneous neural replay account for (a) déjà vu experiences and (b) involuntary autobiographical memories (IAMs).

hippocampal replay, the reactivated representations in the DMN are unable to be integrated into a complete autobiographical memory. Instead, transient activity in the DMN may correspond to reactivation of some stored features or attributes of an autobiographical memory, thereby evoking a vague sense of familiarity to specific features typical of déjà vu experiences. Complete memory retrieval of IAMs and access to their memory contents occur only when spontaneous neural replay occurs in both the hippocampus and DMN (Fig 1b). Reactivation of hippocampal highest-order memory index codes allows for the binding or integration of the various factorized feature representations reactivated in the DMN, resulting in a successful involuntary retrieval of an autobiographical memory with explicit accessibility to its contents.

Further support to our spontaneous neural replay account comes from lesion studies. McCormick, Rosenthal, Miller, and Maguire (2018) found that patients with bilateral hippocampal damage could engage in spontaneous thinking but were unable to report contents of their spontaneous thoughts about their memories. This suggests hippocampus plays a critical role in the complete and successful retrieval of IAMs. Additionally, the earliest/preclinical stages of Alzheimer’s disease (AD) are characterized by accumulation of extracellular β -amyloid primarily in DMN nodes such as PCC and mPFC which produces cell atrophy and disrupted functional connectivity in the DMN (Palmqvist et al., 2017). Such disruption to DMN in early AD has been associated with reductions in the rate of spontaneous thoughts about the past without impairing accessibility to their contents (Kvavilashvili, Niedźwieńska, Gilbert, & Markostamou, 2020; O’Callaghan, Shine, Lewis,

Andrews-Hanna, & Irish, 2015). Taken together, disruption to hippocampus and DMN may impair distinct aspects (accessibility to memory contents vs. frequency of spontaneous recall) of involuntary retrieval of autobiographical memories.

Future studies are warranted to directly dissociate the contribution of hippocampus and DMN in spontaneous memory retrieval. It would be advantageous to examine spontaneous neural replay across cortical DMN nodes and subcortical structure of hippocampus by combining neuroimaging techniques that offer both high temporal and spatial resolutions (e.g., magnetoencephalography). In conclusion, we have implicated spontaneous neural replay in hippocampus and DMN as the neural basis for IAMs while transient activation limited to DMN supports déjà vu experiences. By highlighting the critical role of spontaneous neural replay in involuntary memory processes and how replay in different neural substrates contributes distinctly to involuntary retrieval, we offer a more nuanced neurocognitive perspective on how IAMs and déjà vu experiences differ.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences*, 11(5), 219–227.

- Deuker, L., Olligs, J., Fell, J., Kranz, T. A., Mormann, F., Montag, C., ... Axmacher, N. (2013). Memory consolidation by replay of stimulus-specific neural activity. *Journal of Neuroscience*, 33(49), 19373–19383.
- Dickey, C. W., Verzhbinsky, I. A., Jiang, X., Rosen, B. Q., Kajfez, S., Stedelin, B., ... Halgren, E. (2022). Widespread ripples synchronize human cortical activity during sleep, waking, and memory recall. *Proceedings of the National Academy of Sciences*, 119(28), e2107797119.
- Gillespie, A. K., Maya, D. A. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., ... Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron*, 109(19), 3149–3163.
- Goode, T. D., Tanaka, K. Z., Sahay, A., & McHugh, T. J. (2020). An integrated index: Engrams, place cells, and hippocampal memory. *Neuron*, 107(5), 805–820.
- Higgins, C., Liu, Y., Vidaurre, D., Kurth-Nelson, Z., Dolan, R., Behrens, T., & Woolrich, M. (2021). Replay bursts in humans coincide with activation of the default mode and parietal alpha networks. *Neuron*, 109(5), 882–893.
- Kaefer, K., Stella, F., McNaughton, B. L., & Battaglia, F. P. (2022). Replay, the default mode network and the cascaded memory systems model. *Nature Reviews Neuroscience*, 23(10), 628–640.
- Kvavilashvili, L., Niedźwieńska, A., Gilbert, S. J., & Markostamou, I. (2020). Deficits in spontaneous cognition as an early marker of Alzheimer's disease. *Trends in Cognitive Sciences*, 24(4), 285–301.
- Liu, Y., Nour, M. M., Schuck, N. W., Behrens, T. E., & Dolan, R. J. (2022). Decoding cognition from spontaneous neural activity. *Nature Reviews Neuroscience*, 23(4), 204–214.
- McCormick, C., Rosenthal, C. R., Miller, T. D., & Maguire, E. A. (2018). Mind-wandering in people with hippocampal damage. *Journal of Neuroscience*, 38(11), 2745–2754.
- O'Callaghan, C., Shine, J. M., Lewis, S. J., Andrews-Hanna, J. R., & Irish, M. (2015). Shaped by our thoughts – a new task to assess spontaneous cognition and its associated neural correlates in the default network. *Brain and Cognition*, 93, 1–10.
- Palmqvist, S., Schöll, M., Strandberg, O., Mattsson, N., Stomrud, E., Zetterberg, H., ... Hansson, O. (2017). Earliest accumulation of β -amyloid occurs within the default-mode network and concurrently affects brain connectivity. *Nature Communications*, 8(1), 1–13.
- Philippi, C. L., Tranel, D., Duff, M., & Rudrauf, D. (2015). Damage to the default mode network disrupts autobiographical memory retrieval. *Social Cognitive and Affective Neuroscience*, 10(3), 318–326.
- Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, 364(6447).
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, 20(1), 11.
- Simons, J. S., Ritchey, M., & Fernyhough, C. (2022). Brain mechanisms underlying the subjective experience of remembering. *Annual Review of Psychology*, 73, 159–186.
- Vaz, A. P., Wittig Jr., J. H., Inati, S. K., & Zaghoul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. *Science*, 367(6482), 1131–1134.

Evolutionary mismatch and anomalies in the memory system

Walter Veit^a  and Heather Browning^b 

^aDepartment of Philosophy, University of Bristol, Bristol, UK and ^bDepartment of Philosophy, University of Southampton, Southampton, UK
wvveit@gmail.com
DrHeatherBrowning@gmail.com
<https://walterveit.com/>
<https://www.heatherbrowning.net>

doi:10.1017/S0140525X23000183, e381

Abstract

In order to understand involuntary autobiographical memories and déjà vu experiences, we argue that it is important to take an evolutionary medicine perspective. Here, we propose that these memory anomalies can be understood as the outcomes of an inevitable design trade-off between type I and type II errors in memory processing.

Barzykowski and Moulin (B&M) offer an excellent synthesis of a wealth of empirical data from a variety of disciplines to simultaneously explain two spontaneous phenomena of the memory

system that have so far eluded satisfying explanation: Involuntary autobiographical memories (IAMS) and déjà vu experiences. Whereas the former are invasive recollections of the personal past, the latter constitute brief experiences of familiarity while simultaneously knowing that the familiarity is false. Like the authors, we think that any theory or account of memory retrieval should account for the apparently pathological or dysfunctional anomalies of the memory system. Furthermore, we find their account extremely compelling, particularly as it places IAMS and déjà vu on a continuum with both involving what Barzykowski and Moulin describe as involuntary cognitions, and our commentary is not at all intended as a criticism of their hypothesis.

Instead, we hope to make use of this opportunity to further advance their proposal by focusing on the ultimate or evolutionary explanation for the phenomena. B&M primarily concentrate on the mechanisms and triggers; providing a convincing proximate explanation for IAMS and déjà vu. However, only in their conclusion do they briefly consider an evolutionary function, suggesting that these memory anomalies can be seen as “the result of a continuously active memory system that automatically and rapidly scans the environment for matching representations”; a suggestion which we would like to expand on. We think there is much promise in the idea that the brain is continually and rapidly scanning the environment for opportune information and attempting to match this to relevant stored representations, a process that sometimes intrudes into conscious awareness. Indeed, it would allow us to explain both IAMS and déjà vu experiences as evolutionary mismatches, phenomena that have received much attention in the evolutionary medicine literature (see Manus 2018; Stearns, 2012; Veit & Browning, 2021). Since our modern environments contain many more stimuli than the ancestral ones in which our memory system evolved, it should not be at all surprising that there can be frequent instances of misfiring, especially when – as in the cases of anomalies such as IAMS and déjà vu experiences – there does not appear to be an immediate fitness cost.

From an evolutionary perspective, there could thus be a straightforward design trade-off in building a costly memory system that has to pay off for the organism to be functional. Since organisms stand to benefit greatly from having pertinent information raised to conscious awareness, while false positives in the form of déjà vu experiences and IAMS have little cost in terms of fitness, it makes sense that evolution would favour the avoidance of type II errors (false negatives such as failing to remember important familiar situations) over type I errors (false positives such as mistakenly thinking that a place is familiar). While it may seem intuitive to think that healthy forms of cognition should not have any anomalies of this sort, to do so would be a failure to recognize that these error rates are inversely related to each other, and thus cannot both be minimized at the same time. There are trade-offs and it is plausible that evolution designed the memory system to prioritize the minimization of type II errors.

Nevertheless, even if there is such a trade-off, that does of course not mean that type II errors are *always* to be preferred over type I errors. As B&M themselves acknowledge, feelings of familiarity can be *pathologically overactive*, where inputs are repeatedly accompanied by feelings of familiarity that Moulin (2013) describes as recollective confabulation. What should we make of these cases that are akin to a permanent déjà vu? From an evolutionary medicine perspective, we should not at all be surprised that neuropathological cases can be found, in which these evolutionary trade-offs are handled in a dysfunctional manner. Indeed, these cases may provide us with the best source of

evidence for understanding how natural selection has dealt with trade-offs in “designing” the human memory system. Importantly, if we want to understand such pathological cases of the mind, it is important to put evolutionary thinking centre stage, since it is only from a Darwinian design stance that we will be able to understand what makes apparent anomalies of the memory system pathological (or for that matter, healthy) (Veit & Browning, 2023). After all, it is precisely in asking for the costs and benefits of different kinds of type I and type II errors that we can begin to understand the memory system as a teleonomic system designed to maximize the fitness of organisms.

Finally, we would like to again reiterate that we believe there to be much promise in the account of B&M. Nevertheless, in order to advance their proposal, we propose that there would be a benefit in studying the anomalies of the memory system framed as type I and type II errors. This could lead to more precise hypotheses that could in turn be tested. Indeed, we may even be able to derive computational models and simulations in order to study these trade-offs and under which environmental conditions there may be fitness advantages to investing in the avoidance of one error over the other. While we have been sceptical of very ambitious attempts to model all phenomena of the mind in terms of free energy minimization or predictive error minimization (Veit & Browning, 2022), this may be a good case for where this framework could legitimately help us to further our understanding of how the brain deals with errors in the memory system and why some errors are evolved features of the architecture of our minds.

Financial support. This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement number 101018533).

Competing interest. None.

References

- Manus, M. B. (2018). Evolutionary mismatch. *Evolution, Medicine, and Public Health*, 2018(1), 190–191.
- Moulin, C. J. (2013). Disordered recognition memory: Recollective confabulation. *Cortex*, 49(6), 1541–1552.
- Stearns, S. C. (2012). Evolutionary medicine: Its scope, interest and potential. *Proceedings of the Royal Society B: Biological Sciences*, 279(1746), 4305–4321.
- Veit, W., & Browning, H. (2021). Developmental programming, evolution, and animal welfare: A case for evolutionary veterinary science. *Journal of Applied Animal Welfare Science*, 1–13. <https://doi.org/10.1080/10888705.2021.2014838>
- Veit, W., & Browning, H. (2022). Life, mind, agency: Why Markov blankets fail the test of evolution. *Behavioral and Brain Sciences*, 45, E214. <https://doi.org/10.1017/S0140525X22000115>
- Veit, W., & Browning, H. (2023). Hominin life history, pathological complexity, and the evolution of anxiety. *Behavioral and Brain Sciences*, 46, e79. doi: [10.1017/S0140525X22001923](https://doi.org/10.1017/S0140525X22001923)

Involuntary memory signals in the medial temporal lobe

Haopei Yang^a, Chris B. Martin^b and Stefan Köhler^c 

^aGraduate Program in Neuroscience, Western University, London, ON, Canada;

^bDepartment of Psychology, Florida State University, Tallahassee, FL, USA and

^cDepartment of Psychology, Western University, London, ON, Canada

hyang336@uwo.ca

cmartin@psy.fsu.edu

stefank@uwo.ca

<https://martinmemorylab.com/>

<https://kohlermemorylab.org/>

doi:10.1017/S0140525X23000067, e382

Abstract

We highlight recent progress in neuroimaging and neuropsychological research on memory mechanisms in the medial temporal lobe that speaks to the involuntary nature of memory retrieval processes. We suggest that evidence from these studies supports Barzykowski and Moulin’s proposal that memory signals involved in experiences of familiarity and déjà vu can be generated in the absence of retrieval intentionality.

We commend Barzykowski and Moulin (B&M) on presenting a theoretical model that considers links between mnemonic experiences that are typically not discussed together in the cognitive-psychology and cognitive-neuroscience literature. We agree with the emphasis on the shared involuntary nature of the memory experiences covered in the model, and the proposed central role that familiarity plays, as summarized in the schematic in Figure 1. At the same time, we feel that this model could be further developed at the mechanistic level through consideration of the functional properties of perirhinal cortex (PrC) in the medial temporal lobe; it is the brain region that has been most closely linked to familiarity assessment in extant research. To this end, we highlight several recent findings, not covered in the target article, that speak to the involuntary mode in which familiarity signals can arise.

PrC has been implicated in familiarity across many functional neuroimaging studies conducted in neurologically healthy participants and in research in individuals with brain damage (Köhler & Martin, 2020; Montaldi & Kafkas, 2022). Although the majority of reports comprising this literature establish a link between PrC and item familiarity in experimental contexts that involve intentional retrieval (e.g., Montaldi, Spencer, Roberts, & Mayes, 2006), a growing body of evidence suggests that this structure also supports involuntary effects of familiarity or fluency that can be observed in the absence of retrieval intentionality. For example, activity in PrC tracks experimental exposure history for verbal and non-verbal stimuli in tasks that show behavioural priming effects (Voss, Hauner, & Paller, 2009; Wang, Ranganath, & Yonelinas, 2014; Yang, McRae, & Köhler, 2023). In related functional magnetic resonance imaging (fMRI) work, we have recently shown that PrC not only tracks recent laboratory exposure, but also degree of judged lifetime exposure to object concepts outside the lab (Duke, Martin, Bowles, McRae, & Köhler, 2017; Yang et al., 2023). Specifically, using a paradigm in which participants made judgements that either required or did not require consideration of lifetime familiarity, we found that fMRI BOLD activity in PrC tracked this memory characteristic regardless of retrieval intentionality. Moreover, observed behavioural performance was also sensitive to lifetime familiarity under conditions in which the latter was task irrelevant (Yang et al., 2023). Overall, such evidence points to PrC as a structure that may support the cue-induced familiarity that is central to involuntary memory retrieval in B&M’s model.

A role for PrC in involuntary memory has also been revealed in research conducted in neurological patients who experience déjà vu

in the context of seizures with medial temporal-lobe origin (i.e., ictal déjà vu). These experiences have several phenomenological characteristics that mirror those highlighted by B&M for spontaneously occurring déjà vu in neurotypical individuals. At the core, ictal déjà vu is also characterized by an involuntarily arising feeling of familiarity that conflicts with a metacognitive awareness of situational novelty. Although ictal déjà vu in temporal-lobe epilepsy is a transient phenomenon, the eliciting seizure activity is tied to lasting structural and functional brain abnormalities. These abnormalities provide a window towards understanding what brain regions may be involved in triggering the familiarity experience that is central to déjà vu. Indeed, we have reported that ictal déjà vu is associated with quantifiable abnormal brain structure in PrC that also goes hand in hand with persistent familiarity-based memory impairments on experimental tasks of recognition memory (Martin et al., 2021).

While we suggest that the neural familiarity signals that are central to the involuntary familiarity and déjà vu experiences described in B&M's model are closely tied to PrC functions, we would not argue that activity in this region alone is *sufficient* to generate these mnemonic experiences. Rather, we agree with the authors' proposal that automatically generated familiarity signals may only serve as an initial trigger; depending on the outcome of downstream processes that include metacognitive evaluation of situational novelty, they may or may not lead to memory experiences of familiarity or déjà vu. As such, we do not want to leave the impression that we disagree with the authors' point of emphasis on large-scale brain networks, such as the default mode network, when attempting to explain involuntary memory retrieval. However, given the increasing appreciation of the complexity of memory-retrieval cascades in the cognitive neuroscience literature, we think that linking specific brain regions to specific component processes holds unique promise for further mechanistic and theoretical development of the present model.


Financial support. This research was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (Canada) to S. K.

Competing interest. None.

References

- Duke, D., Martin, C. B., Bowles, B., McRae, K., & Köhler, S. (2017). Perirhinal cortex tracks degree of recent as well as cumulative lifetime experience with object concepts. *Cortex*, 89, 61–70. <https://doi.org/10.1016/j.cortex.2017.01.015>
- Köhler, S., & Martin, C. B. (2020). Familiarity impairments after anterior temporal-lobe resection with hippocampal sparing: Lessons learned from case NB. *Neuropsychologia*, 138, 107339. <https://doi.org/10.1016/j.neuropsychologia.2020.107339>
- Martin, C. B., Mirsattari, S. M., Pruessner, J. C., Burneo, J. G., Hayman-Abello, B., & Köhler, S. (2021). Relationship between déjà vu experiences and recognition-memory impairments in temporal-lobe epilepsy. *Memory*, 29, 884–894. <https://doi.org/10.1080/09658211.2019.1643891>
- Montaldi, D., & Kafkas, A. (2022). Neural mechanisms of familiarity. *PsyArXiv*. <https://doi.org/10.31234/osf.io/zfq5s>
- Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus*, 16(5), 504–520. <https://doi.org/10.1002/hipo.20178>
- Voss, J. L., Hauner, K. K. Y., & Paller, K. A. (2009). Establishing a relationship between activity reduction in human perirhinal cortex and priming. *Hippocampus*, 19(9), 773–778. <https://doi.org/10.1002/hipo.20608>
- Wang, W.-C., Ranganath, C., & Yonelinas, A. P. (2014). Activity reductions in perirhinal cortex predict conceptual priming and familiarity-based recognition. *Neuropsychologia*, 52, 19–26. <https://doi.org/10.1016/j.neuropsychologia.2013.10.006>
- Yang, H., McRae, K., & Köhler, S. (2023). Perirhinal cortex automatically tracks multiple types of familiarity regardless of task-relevance. *Neuropsychologia*, 187, 108600. <https://doi.org/10.1016/j.neuropsychologia.2023.108600>

Oh it's me again: Déjà vu, the brain, and self-awareness

Samantha Zorns, Claudia Sierzputowski,
Matthew Pardillo and Julian Paul Keenan 

Cognitive Neuroimaging Laboratory, Montclair State University, Montclair, NJ, USA

zornss1@montclair.edu

sierzputowsc1@montclair.edu

pardillom1@montclair.edu

keenanj@montclair.edu

www.cognitiveneuroimaginglab.com

doi:10.1017/S0140525X23000201, e383

Abstract

Déjà vu and involuntary autobiographical memories (IAMs) are differentiated by a number of factors including metacognition. In contrast to IAMs, déjà vu activates regions associated with self-awareness including the right dorsolateral prefrontal cortex.

Barzykowski and Moulin (B&M) provide us with an overview of two related processes and provide significant detail of the psychological mechanisms. While the authors touch on the underlying neural correlates (e.g., cortical midline structures, medial-temporal), more detail as to the proximate brain mechanisms would further convince readers of their claims. Further, but not addressed here, the authors leave the readers wanting a deeper evolutionary, ultimate perspective and we look forward to papers that address evolutionary, comparative, and genetic examinations of these two memory phenomena.

As elucidated by B&M, oddities of human cognition often provide significant insight into the mechanistic functions of the brain, which in turn provides knowledge about the phenomena themselves. Both involuntary autobiographical memories (IAMs) and déjà vu are two such examples. IAMs are spontaneously recalled without conscious effort (Berntsen, 2021), are often triggered by elements of the environment that correspond to a previous event (Berntsen, 2009), and are typically unpleasant, and as such, they are one of the main symptoms of posttraumatic stress disorder (PTSD). B&M note that IAMs are a feature of normal functioning cognitive processes, and it is the frequency and intensity (rather than their existence) that warrants a PTSD diagnosis (Rasmussen & Berntsen, 2011). In contrast to IAMs, déjà vu is categorized as any erroneous subjective perception that a currently experienced event is similar to an indeterminate past occurrence (Sno & Linszen, 1990). Episodes of déjà vu originate from a memory error that is often influenced by feelings of fluency and familiarity, thus affecting memory retrieval (Illman, Butler, Souchay, & Moulin, 2012).

However, the authors may wish to consider a bigger role for “the self” in these processes, as evidence from neuroscience guides us to think higher order “meta” processes may distinguish these memory deviations. Metacognition, as defined here, refers to the capacity to keep track of and regulate one's own cognitive processes (Fleur, Bredeweg, & van den Bos, 2021). The link between déjà vu and the sensation of foreboding may be explained by the

metacognitive bias initiated by the experience itself (Cleary & Claxton, 2018). In brief, IAMs typically have a reduction of metacognitive processes when compared to déjà vu which involves a sense of seeing oneself as an agent in the experience. We derive this notion from neurological studies in both normals and patients.

In terms of neural correlates, IAMs typically activate the left angular gyrus, posterior midline, visual cortex, hippocampus, and parahippocampal cortex (Berntsen, 2021). Contrary to IAMs, the medial prefrontal cortex is engaged during déjà vu. As a further distinction between IAMs and déjà vu, and likely more important in terms of the behavioral differences, the right hemisphere appears active during bouts of déjà vu (Vlasov, Chervyakov, & Gnezditskii, 2013). The right frontal cortex and temporal parietal junction are active in normals during tasks that involve self-awareness (Morita et al., 2018) and disruption of the right hemisphere disrupts self-awareness (Keenan, Nelson, O'Connor, & Pascual-Leone, 2001).

Déjà vu is but one of many perturbations that involves the right frontal/temporal regions and the self. While déjà vu involves a brief narrative clip, significant neural damage can lead to truly devastating disorders. The occurrence of hallucinations in which a location has been duplicated, or exists in two different places, is known as reduplicative paramnesia, which is a rare illness seen in neurological patients (Borghesani, Monti, Fortis, & Miceli, 2019). An example of reduplicative paramnesia involved a 70-year-old woman, who presented to her local emergency room, and was found to have a right temporal meningioma. In the days prior to her craniotomy, she had been able to recount “her two houses.” She was aware she owned an apartment in Rome, yet in the previous month, her single apartment had “become” two interchangeable homes, with identical floors, addresses, and furniture (Gerace & Blundo, 2013). Capgras syndrome, also referred to as the delusion of doubles, is defined by the patient’s mistaken conviction that an exact duplicate has taken the place of a person of significance (Shah, Jain, & Wadhwa, 2022). Within the disorder, the delusion typically involves the afflicted person assuming their close friend or relative has disappeared. The impersonator remains without name and identity, and is viewed pejoratively (Barrelle & Luauté, 2018). Fregoli syndrome is defined as the irrational mistaken identification of known individuals posing as others, as well as unfamiliar strangers and/or casual acquaintances identified as a different but known individual (Langdon, Connaughton, & Coltheart, 2014). The disorder has frequently been linked to Capgras syndrome in the past; however, the two have distinct defining features (Mojtabai, 1994).

There are numerous other, similar disorders of misidentification of self and other (Feinberg & Keenan, 2005). We believe, like B&M, the distinction between IAMs and déjà vu provide insights into human cognition that reach beyond the phenomena themselves. Here we hope to have elucidated a number of proximate reasons for such a distinction. However, we believe that the role of the self may play a larger role than posited by B&M. Neuroimaging demonstrates that déjà vu activates regions associated with self-awareness and we believe this is one of the cases where neuroscience can help to clarify a psychological debate.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

- Barrelle, A., & Luauté, J. P. (2018). Capgras syndrome and other delusional misidentification syndromes. *Frontiers of Neurology and Neuroscience*, 42, 35–43. doi:10.1155/000475680
- Berntsen, D. (2009). *Involuntary autobiographical memories: An introduction to the unbidden past*. Cambridge University Press.
- Berntsen, D. (2021). Involuntary autobiographical memories and their relation to other forms of spontaneous thoughts. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 376(1817), 20190693. doi:10.1098/rstb.2019.0693
- Borghesani, V., Monti, A., Fortis, P., & Miceli, G. (2019). Reduplicative paramnesia for places: A comprehensive review of the literature and a new case report. *Clinical Neurology and Neurosurgery*, 181, 7–20. doi:10.1016/j.clineuro.2019.03.022
- Cleary, A. M., & Claxton, A. B. (2018). Déjà vu: An illusion of prediction. *Psychological Science*, 29(4), 635–644. doi:10.1177/0956797617743018
- Feinberg, T. E., & Keenan, J. P. (2005). Where in the brain is the self? *Consciousness and Cognition*, 14(4), 661–678. doi:10.1016/j.concog.2005.01.002
- Fleur, D. S., Bredeweg, B., & van den Bos, W. (2021). Metacognition: Ideas and insights from neuro- and educational sciences. *NPJ Science of Learning*, 6(1), 13. doi:10.1038/s41539-021-00089-5
- Gerace, C., & Blundo, C. (2013). Reduplicative paramnesia: Not only one. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 25(3), E16–E18. doi:10.1176/appi.neuropsych.12030072
- Illman, N. A., Butler, C. R., Souchay, C., & Moulin, C. J. (2012). Déjà experiences in temporal lobe epilepsy. *Epilepsy Research and Treatment*, 2012, 539567. doi:10.1155/2012/539567
- Keenan, J. P., Nelson, A., O'Connor, M., & Pascual-Leone, A. (2001). Self-recognition and the right hemisphere. *Nature*, 409(6818), 305. doi:10.1038/35053167
- Langdon, R., Connaughton, E., & Coltheart, M. (2014). The Fregoli delusion: A disorder of person identification and tracking. *Topics in Cognitive Science*, 6(4), 615–631. doi:10.1111/tops.12108
- Mojtabai, R. (1994). Fregoli syndrome. *Australian and New Zealand Journal of Psychiatry*, 28(3), 458–462. doi:10.3109/00048679409075874
- Morita, T., Saito, D. N., Ban, M., Shimada, K., Okamoto, Y., Kosaka, H., ... Naito, E. (2018). Self-face recognition begins to share active region in right inferior parietal lobule with proprioceptive illusion during adolescence. *Cerebral Cortex*, 28(4), 1532–1548. doi:10.1093/cercor/bhy027
- Rasmussen, A. S., & Berntsen, D. (2011). The unpredictable past: Spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition*, 20(4), 1842–1846. doi:10.1016/j.concog.2011.07.010
- Shah, K., Jain, S. B., & Wadhwa, R. (2022). Capgras syndrome. In StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.
- Sno, H. N., & Linszen, D. H. (1990). The déjà vu experience: Remembrance of things past? *American Journal of Psychiatry*, 147(12), 1587–1595. doi:10.1176/ajp.147.12.1587
- Vertes, R. P. (2006). Interactions among the medial prefrontal cortex, hippocampus and midline thalamus in emotional and cognitive processing in the rat. *Neuroscience*, 142(1), 1–20. doi:10.1016/j.neuroscience.2006.06.027
- Vlasov, P. N., Chervyakov, A. V., & Gnezditskii, V. V. (2013). Déjà vu phenomenon-related EEG pattern. Case report. *Epilepsy & Behavior Case Reports*, 1, 136–141. doi:10.1016/j.ebcr.2013.08.001

Authors' Response

Further advancing theories of retrieval of the personal past

Krystian Barzykowski^{a*} and Chris J. A. Moulin^{b,c*}

^aApplied Memory Research Laboratory, Faculty of Philosophy, Institute of Psychology, Jagiellonian University, Kraków, Poland; ^bLaboratoire de Psychologie et Neurocognition, Université Grenoble Alpes, Grenoble, France and ^cInstitut Universitaire de France
krystian.barzykowski@uj.edu.pl

doi:10.1017/S0140525X23002765, e384

*The authors share first authorship.

Abstract

In our target article, we presented the idea that involuntary autobiographical memories (IAMs) and déjà vu may both be based on the same retrieval processes. Our core claim was thus straightforward: Both can be described as “involuntary” or spontaneous cognitions, where IAMs deliver content and déjà vu delivers only the feeling of retrieval. Our proposal resulted in 27 commentaries covering a broad range of perspectives and approaches. The majority of them have not only amplified our key arguments but also pushed our ideas further by offering extensions, refinements, discussing possible implications and providing additional empirical, neuroscientific and clinical support. The discussion launched by the commentaries proves to us the importance of bringing IAMs and déjà vu into mainstream discussions of memory retrieval processes.

R1. Introduction

The advancement of knowledge is, among other things, about putting forward ideas for discussion. Our main goal was to launch a discussion about the possibility that two seemingly different phenomena, involuntary autobiographical memories (IAMs) and déjà vu, may have the same neurocognitive origin and point to common retrieval processes. Our main claim was: IAMs and déjà vu are two forms of spontaneous experiences derived from autobiographical memory processing. If nothing else, we wanted to bring déjà vu and involuntary memory phenomena into mainstream discussions of memory retrieval processes.

We received 27 commentaries. Let us thank the commentators for their insightful comments, suggestions and thoughts. They allowed us to reconsider new perspectives that ultimately advance the theoretical proposal. On occasion, as developed below, we agree to disagree, especially with those which argue against studying IAMs and déjà vu together.

At its core, our proposal was about the interaction between the environment and internal representations of the personal past, and two key theoretical concepts underpin our thinking: The roles of cues in retrieval and familiarity. The role of cues was mentioned in many commentaries. For instance, **Cleary, Poulos, & Mills (Cleary et al.)** propose a mechanism of *cue familiarity detection*. Depending on the strength of the cue-memory overlap, the resulting familiarity “signal” may pass the critical threshold directing attention inward to initiate a memory search. This notion advances not only our framework but theories of autobiographical retrieval in general. Familiarity and neuroscientific mechanisms of retrieval also make up a large part of the commentaries received.

We also made the claim that to understand déjà vu and IAMs, the processes involved in memory retrieval and in conscious experience appear to be separable, something emphasised in **Schwartz & Pournaghдали’s** commentary. We have recently expanded on the phenomenology of retrieval processes (Moulin, Carreras, & Barzykowski, 2023), including discussion of the tip-of-the-tongue experience (which was largely cut from the target article due to space constraints). Our point is that the brain is constantly trying to make sense of its inputs: In memory, we propose that the phenomenology of retrieval points to there being an attributional system which oversees the function of the temporal lobe. To organise our responses around these key concepts, we discuss conceptual

issues, notably the role of cues in retrieval and the definition of familiarity, and debating the very definition and characteristics of the phenomena. Finally, we turn our attention to the neuroscientific aspects of our conceptualisation of déjà vu and IAMs in a second section before attempting to defend our continuum in the conclusions.

R2. How to conceptualise déjà vu and involuntary memories?

Kvavilashvili & Markostamou focus on cues as a means of explaining two issues: Temporal orientation and the frequency of occurrence. They propose that IAMs seem to be triggered by an easily identifiable cue, whereas déjà vu seems to be more dependent on less specific cues (e.g., the situation as a whole, contextual cues, complex constellation of cues, situational gist properties). Non-specific cues may then activate not the fragments of sensory-perceptual experiences (that are part of the autobiographical memory knowledge system) but rather the representations of previously experienced scenes (that are part of long-term perceptual-representation system). We welcome this suggestion and agree that déjà vu may emerge through several not mutually exclusive paths, and this may be one possible route. However, we consider that activating information from any long-term perceptual-representation system would not explain the intense feeling of self. There is a difference between saying “this reminds me of something” and “this feels like I have been here before.” Therefore, there is space to further advance this idea by explaining how such activation leads to having a “personal past” component.

Robins & Afifi argue that our unified account depends upon how easily concepts from the study of voluntary memory transfer to involuntary phenomena. We agree that cues may differently be processed in involuntary and voluntary retrieval, but we do not think that these differences significantly complicate our framework. While the authors say that “the cue can be identified prior to and independently of the memory it activates,” this is only true in successful recall of a sought-for memory. Moreover, it seems that this falls foul of the Norman and Bobrow (1979) conundrum which we reproduced as a footnote in the target article. In an experiment, or with a degree of introspection, it is easy for an expert observer to pinpoint a cue. In a naturalistic setting, the definition of what the cue *was* is arguably post-hoc and is as subjective as the contents of retrieval, as Robins & Afifi state: “a cue can only ever be identified in retrospect.”

In voluntary retrieval, we have an explicit idea of what we want to recall, but whether a cue will or will not be successful in retrieving it is not a given. We argue that the same is true for involuntary remembering, as we can never be sure whether something in the environment will or will not trigger a given memory, although on occasion a cue is highly efficient in triggering a given memory, for example, intrusive, recurrent memories triggered by easily identified stimuli. We state that retrieval is an iterative process, and what starts as involuntary retrieval may cross over into voluntary retrieval and vice-versa (see also Barzykowski & Mazzoni, 2022).

Evidently, involuntary retrieval is less predictable, as it is spontaneous and automatic. However, we do not agree with **Robins & Afifi** since we propose that people can be aware of the cues triggering memories prior to a complete retrieval. We take issue with the idea that “there is also no sense to be made of the idea of a cue failing to activate a memory during involuntary retrieval. If nothing is activated, then there is no cue – and also no memory.”

When saying that a cue failed to activate a memory during involuntary retrieval, we do not refer to a situation in which nothing is activated but that something is insufficiently activated. Not having memory content does not mean *per se* that some stages of retrieval processes did not happen at all. We propose that there is a prior cascade of processes resulting in retrieval of content: There are instances where there is a cue, but nothing is retrieved. We do not think this is contentious. Detecting something in the environment as familiar based on an evaluation of fluency does not imply retrieval or the access of content, and yet we are able, we think, to identify something as a cue from this feeling, even if nothing is “retrieved.”

Our characterisation of a continuously active memory system trying to make sense of its inputs necessitates that anything in the environment is potentially a cue. For the experient, there is not some philosophical division of the environment into cues and non-cues, but a series of activations and associations between the internal and external world which may or may not enter consciousness.

We now turn our attention to familiarity. A number of commentators interpret familiarity in a different manner from that conceptualised in the article. At least three conceptualisations of familiarity exist (as noted in a footnote by Whittlesea and Williams [1998, p. 141]):

One is that a person has actually encountered a stimulus (or even one like it) previously. ... A second is that the person has knowledge about a stimulus that permits them to perform appropriately toward an object, without necessarily having an accompanying feeling of having experienced that stimulus previously. ... A third is the subjective feeling of having encountered a stimulus on some previous occasion, whether one actually has or not.

Pan & Carruthers helpfully differentiate two types of familiarity. The first type is phenomenological and arises in response to a retrieval failure. The second is akin to that in the recognition-memory literature referring to “a decision-making process based on feelings of fluency.” These definitional issues are critical, and we emphasise that our framework focuses on phenomenological familiarity. A number of commentators point to this lack of a solid basis on which to construct our argument. **Stendardi, Basu, Treves, & Ciaramelli (Stendardi et al.)** claim that the experience of déjà vu is fundamentally different from other forms of familiarity experience, including the *butcher on the bus* which we use as a reference point, because it is based on a configuration of a complex set of factors, and not a single item. As such, they invoke a framework of compositional memory formation (Ryom, Stendardi, Ciaramelli, & Treves, 2022). They propose that déjà vu is an incomplete memory state (and thus far our accounts do not differ) but that it occurs when “some” familiar items from a compositional memory (a rich representation of complex scene from the personal past) are activated (see their figure) in the neocortex, but where a full memory is not retrieved from the hippocampus.

The critical distinction between their proposal and our original account would be that a multitude of cues aiming to reconstruct a rich configuration of information is critical, whereas we were somewhat neutral as to the number of elements which might make up a cue. We suppose that differentiating between their elaboration and the initial proposal would depend not only on how you define familiarity, but also how you define “cue.” The grain size of the information required to trigger déjà

vu or an involuntary memory is open to debate. Even if the experient in the end identifies any one element as critical, we are sure that cues and representations of the personal past are indeed as complex as **Stendardi et al.** point out. It is of interest that accounts of over-extrapolation of familiarity for specific cues to the whole scene or “single-element familiarity” has long been proposed as a mechanism of how déjà vu is generated (e.g., Leeds, 1944).

Sikorski & Sitek propose a comparison with misidentification syndromes and familiarity. However, the typical conceptualisation of familiarity in misidentification syndromes is as objective and not phenomenological. As such, when they write “While in Capgras syndrome access to memory content is preserved (patient’s proxy face is correctly recognised) [the] feeling of familiarity is missing (the proxy is perceived as an imposter)...” it seems that whether or not this is a genuine first-person “feeling” or an attribution imputed by an observer is a critical question. An important issue therefore to address is whether there is a disorder of phenomenological familiarity in such misidentification syndromes. Whilst we take issue with their implied definition of familiarity, they raise an important point: The role of emotion, proposing that “familiarity has its emotional aspect and cannot be reduced to (meta)-cognitive components.” We think that this is probably more the case in delusional misidentification than it is day-to-day attributions of past experience, but it is something which warrants further investigation. There is sometimes an emotional response to déjà vu and involuntary memories, and moreover, familiarity is invoked in various types of emotional responses (as demonstrated in the mere exposure effect, e.g., Kunst-Wilson & Zajonc, 1980).

A further consideration is that of *gist*, which somewhat overlaps with the idea of a non-specific feeling driven by general information as implied in some conceptualisations of familiarity. **Pan & Carruthers** describe a system with recollection-based identification of the gist of the current scene or event. In agreement, Moulin et al. (2023) propose that a retrieval model is created at the moment of accessing a memory which includes a general model (i.e., gist) of pertinent information from various stores including a metacognitive template of what can be expected from retrieval. In the target article, we did not explicitly discuss the processing of gist information, but this general sense of a scene or event is something captured in the Gestalt account of déjà vu (e.g., Cleary, 2008). Clearly, when cues are not strong enough to sufficiently activate a given memory, there still may be sufficient activation for a gist identification.

With these comments about gist as a specific mechanism in retrieval, we see similar conceptual issues about retrieval schema and the interaction with semantic structures identified in the commentaries by **Irish, Renault & Debrulle** and **Sikorski & Sitek**. Renault refers to *jamais vu*, which has been described as the opposite of déjà vu, whereby an objectively familiar stimulus feels as being unfamiliar (Brown, 2003). It seems to be even less frequent than déjà vu in daily life, although it proves relatively easy to reproduce in the laboratory (Moulin et al., 2021). It can be provoked by repeatedly writing a word until it loses its meaning, and so at the time of writing, it seems to be more of a semantic phenomenon than one about the personal past, and as such it escapes our interest here. However, as **Renault & Debrulle** point out, there is a need to better consider semantic networks and schema more generally in our thinking: Either as a repeated context, a gist representation or even the use of personal semantics to act as a plausibility check.

Indeed, **Morales-Torres & De Brigard** raise the possibility of a relationship between déjà vu and “nonbelieved memories.” If we organise our concepts around two dimensions: (a) access to content (full/absent) and (b) feeling of familiarity (minimal/maximal), we think that non-believed memories differ from déjà vu in the access to content dimension. While déjà vu lacks content, leaving us only with a feeling of familiarity, the non-believed memory has content. Importantly, both are believed to be untrue. In this regard, we agree that non-believed memories – also an unusual and infrequent phenomenon – fall out of our way of thinking about content and feelings (see Moulin et al., 2023, for an explanation of this idea), but it is possible that semantic structures play a greater role in the identification of implausibility in non-believed memories.

To best tackle these issues (definitions of familiarity, involvement of gist and semantic retrieval), we should draw on the neuroscientific literature, and restrict our thinking to known mechanisms of familiarity. Recently, the literature has converged on two types of familiarity, both of which can accommodate our framework, echoing the **Pan & Carruthers** commentary: One which is an attribution based on fluency and the other which derives from perceptual or conceptual overlap (Montaldi & Kafkas, 2022). We explore in more detail these ideas with reference to the brain (see below).

A number of commentaries extend our work beyond memory systems. Our focus was on memory, but it was not our intention to imply that *only* memory processes are important for understanding IAMs and déjà vu, since a number of other cognitive operations come into play in our characterisation of a constantly active process surveying the environment. We are thus in agreement with **Vannucci & Hańczakowski** who propose that attentional factors (e.g., monitoring the stream of consciousness) may affect the awareness threshold. We also find their idea appealing that attention may be a “switch” mechanism responsible for whether common retrieval processes may result in such different states.” Similarly, attentional factors may contribute to the differences in phenomenological characteristics between, for instance, involuntary and voluntary memories (see Barzykowski, Niedźwieńska, & Mazzoni, 2019; Barzykowski & Staugaard, 2018). We are not sure whether attention plays a leading role but we surely agree that attention has an important part to play in both the pre-retrieval stage (i.e., how either peripheral or central/focal potential cues are processed) and the post-retrieval stage (i.e., how it changes the threshold of awareness).

Other commentators consider the very definition of the phenomena being discussed, and given their subjective nature, it is unsurprising that commentators develop, extend and contest our description of these phenomena, especially déjà vu. Some of these commentaries draw out differences between déjà vu and involuntary memory using specific characteristics. One such characteristic was the frequency of the experiences. **Kvavilashvili & Markostamou** invite us to wonder “why déjà vu experiences are so rare compared to the frequent occurrences of IAMs.” In preliminary findings (Zareen, Ahraf, Barzykowski, & Moulin, in preparation) from a large scale-survey of respondents (mean age = 22) in France ($n = 178$), Pakistan ($n = 370$) and Poland ($n = 312$), we found that déjà vu was experienced fewer times per week than IAMs (median values respectively 0 and 1–2 times) and fewer times per year (median values respectively, 3–5 times and 11–20 times).¹

Why would there be these differences in frequency of occurrence? One possibility is that there are different types of IAMs.

Curot, Servais, & Barbeau (Curot et al.) remind us of this issue. They explicitly ask whether there are different types of IAMs with eloquent reference to Marcel Proust’s madeleine. We find this extremely relevant, and existing research can contribute to this discussion. According to Mace (2004) while 30% of all memories reported were triggered by sensory/perceptual cues, only 3% of them were triggered by Proustian-type sensory cues (i.e., tastes or smells). Curot et al. refer to a study by Ball and Little (2006) showing that IAMs retrieved without easily identified cues were a minority. Curot et al. thus argue that these types of IAMs and déjà vu actually share similar prevalence. First, this clearly shows that, while IAMs may be frequently triggered by cues, corroborating their relative high frequency, we overlook IAMs that were retrieved without any identifiable cue or by a “rare” type of cue. As such, déjà vu and IAMs may be more similar in their frequency (also phenomenological characteristics like surprise, or strength) when we look at IAMs triggered in similar situations. Therefore, we should look at IAMs triggered by certain types of cues to make an adequate comparison with déjà vu. Our focus on the cues and different types of IAMs may increase not only their conceptual but also empirical proximity.

Putting aside these issues, let us explore why déjà vu does not happen more frequently. This is exactly the question that has often been asked in IAM research, which we identified in the target article: *Why are we not constantly flooded by IAMs in daily life?* It seems to us that the frequency question pertains then to both déjà vu and IAMs. An interesting avenue proposed by **Morales-Torres & De Brigard** is the possibility to look at the extent to which “novelty is proportional to the mismatch between the stimulus and its context” and the prediction error in which “novelty relates to the divergence between a prior likelihood and a posterior probability.” It seems to us that these possibilities may be relatively easily examined in laboratory settings, and although empirical approaches to déjà vu have considered generating conflict (e.g., Urquhart, Sivakumaran, Macfarlane, & O’Connor, 2021), it would also be of interest to manipulate people’s expectations of what they should experience. If you go into an experiment expecting to have déjà vu, it should be *less* likely to occur if surprise is its key phenomenology, which may be what renders it so difficult to produce anything like genuine déjà vu in an experimental situation!

We do not feel obliged to explain the infrequency of déjà vu as compared to IAMs, so long as we have a sensible, theory-based explanation of how it occurs in the memory system. As an analogue, lightning strikes are relatively infrequent but we understand the conditions by which they are produced. When wanting to explain their infrequency, we fall into a circular (but completely reasonable) argument: Lightning strikes are infrequent because the conditions which lead to their occurrence are infrequent. If déjà vu and IAMs are infrequent, then it is because happily, the cognitive system is such that the conditions by which they are produced are also infrequent.

Andonovski & Michaelian go beyond the frequency issue to claim that our proposal fails to account for the strangeness of déjà vu. We wonder why explaining this one phenomenological property would serve as sine qua non to explain déjà vu properly by the reference to typically occurring processes in autobiographical memory? First, the phenomenological strangeness might simply fall out of the frequency issue. It feels strange because it is rare. Second, per analogiam, we do not have concerns about integrating involuntary memories into autobiographical memory research

despite the robust phenomenological differences between involuntary and voluntary memories.

At this stage, it is important to look for mechanistic reasons for similarities and differences. Similar work has been carried out over the years in involuntary and voluntary memory research; an attempt to explain possible mechanisms resulting in phenomenological differences between involuntary and voluntary memories. A recently proposed answer was the threshold idea which formed the core of the target article. Briefly, retrieval favours phenomenologically “juicy” memories, which we can consider along different dimensions: Attentional factors, memory accessibility, retrieval intentionality, attribution and metacognitive concerned processes, etc. There is, at least for us, no reason to argue why this would be different in retrieval of déjà vu and IAMs.

It is possible that the phenomenological strangeness felt in déjà vu is what makes it pass the awareness threshold. It may be, as suggested by Cleary et al. that depending on the strength of the cue–memory overlap, the resulting familiarity signal has to pass the critical threshold directing attention inward to initiate a memory search. In a retrieval failure, it would bring phenomenological familiarity to one’s attention, because it was strange, unexpected or inappropriate. We have accounted for such a possibility in our proposal, even if we did not use the word “strange.” **Bastin** further expanded on this, proposing that an unexpected content retrieval attempt clashes with feelings and sensations resulting from that retrieval failure that cannot be easily understood or explained. This leads to attribution processes becoming “hyperactive” resulting in strong feelings such as strangeness.

Andonovski & Michaelian also conclude that our proposal does not account for the sudden onset of déjà vu. We did not argue against a sudden onset of déjà vu, and so apologise if we gave that impression. However, over the course of experiencing déjà vu and certainly reflecting metacognitively upon it, it can evolve and change in intensity. In a preliminary analysis (again from Zareen et al. as above), we asked participants to categorise their experience using a series of dichotomies, one of which was to ask participants if the experience came on suddenly or gradually over time. Both déjà vu and IAMs were seen as coming on suddenly by the majority of participants (85% for déjà vu and 62% for IAMs). Interestingly, for tip-of-the-tongue experiences, the majority of participants also reported it as sudden (67%), in contrast with Andonovski & Michaelian’s comment. We are not sure therefore that suddenness is a defining feature of déjà vu, and it does not seem in any case to distinguish IAMs and déjà vu.

We can group their criticism together with a number of other points about micro-phenomenology, and state that we welcome further research into factors such as strangeness, surprise and abruptness. The first step is to accept or reject these somewhat subjective descriptions of déjà vu with reference to empirical studies, and consensual, population-based descriptions, as we have done in the Zareen et al. study cited here. Perhaps the strongest contribution to this phenomenology debate comes from **Perrin**, who proposes that we should further consider déjà vécu as a variety of cognitive experience. Déjà vécu can be thought of as a recollection-based recognition memory error, the direct sibling of déjà vu (see Funkhouser, 1995; Moulin, Conway, Thompson, James, & Jones, 2005; Perrin, Moulin, & Sant’Anna, 2023, for definitions), whereby we feel that the present moment is not only familiar, but could lead to something being *remembered*. It seems reasonable to posit that a recollection form of déjà vu exists given the mechanisms we have proposed in our framework, but we were hesitant to add it at

this stage, since empirical work attesting to the existence of the two types and the reliability of the definitions of the familiarity and recollection forms is severely lacking.

Bastin (see also **Kvavilashvili & Markostamou**) propose that IAMs and déjà vu differ in temporality: Déjà vu is retrieved “with some glimpse into the future with the feeling of prescience.” Similarly, **Addis & Szpunar** suggest that our framework may allow us to better understand related phenomena such as simulations of the future. This leads them to an intriguing speculation; namely, that some instances of déjà vu may be a result of not only weakly activated representations of the past, but also weakly activated (by goal-relevant cues) representations of the future, which they call “déjà vu for the future,” something which may be worthy of future examination in surveys: How often does déjà vu involve a glimpse into the future? It seems to us, that whilst interesting, such prescience is not a ubiquitous feature of this experience (for a single case description of prescience in déjà vu, see Curot et al., 2021), and that future research needs to consider the micro-phenomenology of these experiences in line with Addis & Szpunar’s hypothesis.

A further conceptual concern is the question of functionality. **Mace** elaborated further on the idea that “IAMs, like déjà vu and other similar phenomena, may be cognitive failures.” First of all, we would like to clarify that in the target article, we did not want to adjudicate between two opposing views whereby IAMs are or are not functional. These experiences clearly expose some adaptive function as identified by other commentators (e.g., a constantly active system with a bias for making type I errors [see the evolutionary perspective proposed by **Veit & Browning**]). In addition, we propose that the ability to set an appropriate awareness threshold might have been beneficial from evolutionary perspective for preventing being flooded by spontaneous thoughts (for a case study, see Parker, Cahill, & McGaugh, 2006).

In Table 1, we wrote that IAMs “may be a side effect of typical involuntary processing of contextual/environmental cues, IAMs may serve several functions and roles,” and in retrospect, the word “may” should have been highlighted. We believe that IAMs and déjà vu are most likely the by-products of other cognitive processes, and in this way, any one instance of déjà vu or IAM for the individual is not functional. However, this is not to say that in their apparent randomness, they cannot be functional on some occasions – the individual *may* choose to interpret their IAMs (or indeed déjà vu) or as meaningful or as serving a function. Spontaneous cognitions help us to solve a problem creatively, or remember a prospective task, for example. In any case, IAMs are not completely random, as we were at pains to demonstrate in the target article, since there are several processes and mechanisms showing that the likelihood of their occurrence can be predicted and even manipulated (e.g., Barzykowski & Niedźwieńska, 2018; Mace, 2005).

IAMs and déjà vu “can be explained in the absence of a functional account” (**Mace**). Does this mean that we should treat them as a cognitive failure? If we agree that they are a result of a failure, then we somehow treat them as some sort of exception, but they arise because this is how our cognition works: They are neither failures nor triumphs of the cognitive system, since this evaluation can only be made in such terms in a given context. A helpful parallel might be forgetting. If you forget a friend’s birthday, you may describe it as a failure of the cognitive system, but if you forget the hurtful comments made by a colleague the last time you presented your work at a seminar, we might describe it as adaptive.

To sum up this section, we note that some commentaries extend our thinking into domains other than memory (e.g., attention or inhibitory function discussed by **Burns**) and draw our attention to aspects of the two experiences that we did not discuss in detail (e.g., strangeness, temporality) or which beg questions about our notion of a continuum (e.g., differences in frequency between IAMs and déjà vu). We defend our continuum in a final section, but first offer some concrete mechanisms on which to base our thinking, we respond to the commentaries about the brain.

R3. An elaboration of the neuroscientific basis of déjà vu and IAMs

Neuroanatomically, our proposal is built upon the idea that in human memory, there are separable representational and attribution systems. The importance of the attributional system was highlighted by **Bastin**. We are grateful for her endorsement of our framework since our original neuroscientific view was inspired by Bastin et al.'s (2019) integrative memory framework. Their account explains an interaction between core systems and an attributional system, and it seems to us that any neuroscientific account of memory retrieval must encompass higher-order epimemorial processes captured in the attributions we are able to make about fluency, etc., as well as more mechanistic accounts of the core systems. In a similar way, Montaldi and Kafkas (2022, p. 5) discuss familiarity mechanisms in light of two systems: A global matching system and a fluency-attribution system:

Although the theoretical development of these two familiarity mechanisms have been somewhat distinct, they are highly compatible. ... the global-matching mechanism explains how familiarity signals are computed by brain regions within a potential familiarity memory network ... while the fluency-attribution mechanism explains how the feeling of memory is generated and how decision-making contributes to familiarity memory.

When discussing familiarity, we think we should adhere to these definitions. Of note, these global matching and fluency attribution systems map neatly onto the ideas of gist and phenomenological familiarity discussed above.

A number of commentaries specify the medial temporal-lobe mechanisms which may be responsible for the generation of déjà vu. **Addis & Szpunar** identify pattern completion as a critical mechanism involved in involuntary memory generation. They propose that the hippocampus responds to an overlap between a cue and a stored representation, such that when activation surpasses a certain threshold, retrieval is automatic, the pattern is completed, and then a layer of neurons in the hippocampus, CA3, recruits neocortical areas where different elements of the scene are stored. Their account is compelling, since it is a neuroscientific proposal of a mechanism (i.e., pattern completion) which could account for a threshold-like process in retrieval of the personal past. **Yang, Martin, & Köhler (Yang et al.)** provide a useful coupling between intentionality and déjà vu. We had overlooked in our original article the fact that the overwhelming bias in neuroscientific accounts of familiarity is of research into intentional, strategic detection of familiarity. However, our argument hinges upon an automatic and spontaneous activation of familiarity circuitry. Happily for us, Yang et al. specify research findings which point to involuntary detection of familiarity in the MTL. They point out that the perirhinal cortex specifically

is also activated in response to previously seen stimuli in implicit tasks independently of intentionality. Moreover, if we extend their argument further, it is clear that the perirhinal cortex is also activated during encoding and not simply retrieval (see Montaldi & Kafkas, 2022). As such, we might pinpoint the perirhinal cortex as a critical starting point in a broader network for retrieval.

Although other commentaries specify different structures, such as the anterior thalamic nuclei and the retrosplenial cortex (e.g., **Sikorski & Sitek**), and **Zorns, Sierzputowski, Pardillo, & Keenan (Zorns et al.)** highlight that the right hemisphere might be specifically involved, these commentaries lack the specific mechanisms that are involved in the same way as the **Yang et al.** and **Addis & Szpunar** contributions. While Yang et al. seem to emphasise a view of familiarity as found in the perirhinal cortex, Addis & Szpunar focus on complete retrieval and pinpoint the hippocampus, and in particular CA3 as being a critical feature. Together, these commentaries echo the proposal of **Stendardi et al.**, which essentially suggests déjà vu arises from a lack of hippocampal involvement in retrieval.

Varma & Yu elucidate the aspects of the default mode network in involuntary memories and déjà vu, and further refine the neurological characteristics of these phenomena. As with the **Stendardi et al.** and **Addis & Szpunar** commentaries, they see déjà vu as arising when there is a lack of the usual hippocampal activation in memory retrieval. They invoke the cascaded memory systems framework (Kaefer, Stella, McNaughton, & Battaglia, 2022), and propose spontaneous neural replay as the precise neurocognitive mechanism. Spontaneous neural replay is associated with characteristic bursts of neural activity and occurs in both hippocampal and neocortical areas, and it is thus appealing to think of it as underpinning spontaneous phenomena such as involuntary memories and déjà vu. However, the conscious experience of neural replay is not well understood, and we cannot go too far with this way of thinking since many neuroscientific explorations of spontaneous activity (sharp-wave ripples) are based on rodents, and many analogical processes in the human brain occur during sleep (see Kaefer et al., 2022). That is, whereas in the cascaded memory systems framework, there is constant reactivation of neural patterns which aids in the consolidation and maintenance of memories, it seems unlikely that this spontaneous activity is obligatorily raised to consciousness (especially when we are asleep).

To carry this important neuroscientific work further into rich human experiences like déjà vu and involuntary remembering would require more advanced models of rodent memory: As it stands, we are unsure that a rodent would be able to generate the complex metacognitive evaluation of false familiarity that is at the core of the déjà vu experience. Another strategy proposed by **Varma & Yu** is to use neuroscientific methods, and particularly MEG. This approach has much to offer. For instance, we could use simple associative memory tasks and contrast voluntary and involuntary forms of remembering and contrast neural signatures of the two on otherwise identical tasks. **Curot et al.** offer concrete proposals about how to examine the brain networks involved in déjà vu and IAMs using intracranial electrical brain stimulation, and this is perhaps the best place to start, drawing upon classifications of patients' experiences of IAMs and déjà vu.

Several commentators identify the particular form of temporal-lobe degeneration in dementia as an interesting extension to our work, in keeping with **Kvavilashvili, Niedźwieńska, Gilbert, and Markostamou (2020)** who propose that deficits in spontaneous cognition could be used as an early marker of

Alzheimer's disease (AD). The commentaries do not form a concordant set of proposals, and before discussing them we should point out that there is a lack of empirical work on déjà vu in dementia, and relatively little on involuntary memories. It was beyond the scope of the target article to consider too fully the impact of the healthy ageing process on déjà vu and IAMs, but whilst déjà vu frequency decreases with age, involuntary memory experiences seem to stay somewhat constant across the lifespan (Moulin et al., 2014; see also Maillet & Schacter, 2016; Yeung & Fernandes, 2021, for discussions of age differences in spontaneous cognitions and autobiographical memories). As such, perhaps we need to know more about déjà vu in healthy ageing before considering it in pathological ageing, since base rates of the experience may be so low in dementia as to render it impossible to study.

Bastin; Gautier, Bulteau, Chapelet, & El Haj (Gautier et al.); Irish and Zorns et al. all discuss dementia or the delusions sometimes seen in dementia patients. Whilst **Irish** is too generous in her appraisal to say so explicitly, it is very difficult to reconcile our ideas with what is known about dementia, and thus we too would like to see involuntary memories and déjà vu as “testbeds” for theorising in health and disease. **Irish** focuses on the literature on mindwandering in AD concluding that although some authors report fewer instances of mindwandering in AD, experimental studies show that some self-referential mindwandering is possible. **Gautier et al.** propose that autobiographical retrieval in AD can be viewed “under the lens of a ‘déjà vu’ perspective, lacking the richness of contextual and phenomenological information,” seemingly proposing higher levels of déjà vu in AD. **Irish** also considers whether higher levels of déjà vu *might* be expected to occur in AD, but concludes this seems unlikely. She points to a lack of endogenous generation of mental content in AD as likely to explain the lack of déjà vu and involuntary memories in AD, similar to Kvavilashvili et al. (2020).

We agree with **Irish**, but add that people with AD are known to have deficient metacognitive access in episodic memory (Souchay, 2007; Souchay, Isingrini, & Gil, 2002). As such, people with AD lack the metacognitive access to episodic memory with which to generate the conflict at the core of the déjà vu experience. The schema-driven sense of familiarity described by **Irish** would therefore be more likely to lead to ungated false memories, as there would not be sufficient metacognitive and recollective information to correct this familiarity (e.g., a “recall to reject strategy,” Gallo, Sullivan, Daffner, Schacter, & Budson, 2004; see also Souchay & Moulin, 2009, and for a review of false memories in AD, see El Haj, Colombel, Kapogiannis, & Gallouj, 2020). A straightforward extension of this hypothesis would be that in memory-impaired groups more generally, déjà vu would be negatively correlated with false memories.

Zorns et al. compare déjà vu to different forms of delusional misidentification (see also **Sikorski & Sitek**), and particularly reduplicative paramnesia. These delusions (reviewed in Moulin, 2018) have been explicitly compared with déjà vu. For instance, Feinberg and Shapiro (1989, p. 40) describe a form of reduplication where “the patient maintains that his current experiences are a repeat of past experiences.” Reduplicative paramnesia shows a clear overlap with the concept of recollective confabulation, which has been described as like permanent déjà vu (see Moulin, 2013). The delusion arguably arises because the experient attempts to resolve intact factual knowledge of the world with erroneous feelings or inappropriate affective signals. According to Langdon and Coltheart (2000), this calls for a two-factor

account of the delusion: An underlying cognitive deficit (presumably a problem with over-active familiarity) combined with an erroneous interpretation of the familiarity (the inability to reject the familiarity as false). This is different from déjà vu where the experient is aware of the inappropriateness of the familiarity. In sum, whereas we echo Zorns et al.'s call to explore these delusion syndromes with reference to our proposal, and especially to consider the role of the self in IAMs and déjà vu, there are clear differences between delusion and healthy experiences; notably people with reduplicative paramnesia are anosognosic for their overactive familiarity.

R4. Concluding remarks: The continuum between déjà vu and IAMs

A number of commentaries take issue with the notion of a continuum, and we want to avoid any metaphysical discussions about what a continuum is, only so far as to say that déjà vu and involuntary memories must, according to our view, share some common processes and characteristics. Water, ice and steam share an important common characteristic, but they differ according to context, and might be thought of as lying on a continuum. Is that comparable to the type of shared characteristics we have hypothesised to be at the core of déjà vu and IAMs? Probably not. But until more is known about these phenomena, it seems reasonable to us to point to shared processes available “off-the-shelf” in the memory literature to explain how they occur. The extent to which they operate on a continuum rests upon our notion of the “botched” retrieval of an involuntary memory (to use **Stendardi et al.**'s terminology): On the way to having a fully fledged involuntary memory, we may “only” achieve the strange feeling of familiarity which leads to a déjà vu. **Staugaard** argues explicitly that “the relationship between déjà vu and autobiographical memories is not continuous, but more akin to a path diagram.” Although they might share a starting point, at some point, their paths diverge depending on the presence or absence of memory content. In principle, we agree and this briefly summarises the key claim of our proposition.

Our boldest claim is that déjà vu and IAMs belong to the same family since they operate on the same autobiographical memory base. **Staugaard** and **Gülgöz & Ergen** question whether déjà vu should be considered a part of autobiographical memory. We argue that déjà vu and IAMs may, at least in some circumstances, originate from the same common processes, and are built on autobiographical oriented information (Would we experience déjà vu without having access to our personal past?). But does this make déjà vu an autobiographical memory? We agree with **Staugaard** that this would cause a serious problem for the conventional definition of autobiographical memory. Thus, as long as the instance of déjà vu is embedded within and occurs in the context of one's personal past, we consider déjà vu as an autobiographical memory-affiliated phenomena: IAMs and déjà vu share cognitive mechanisms but at some point their paths diverge. We agree with **Staugaard** that this may happen once the memory activation reaches consciousness. Thus, we agree to disagree with **Gülgöz & Ergen** saying that “clustering these constructs can be counter-productive for research.” We believe just the opposite.

To conclude, we are indebted to our peers for their comments on our work, and they have raised many challenges for empirical and theoretical developments. Some issues, such as the very notion of a continuum, are ideas which are not central to our argument. We can let that idea go as long as we can keep the

idea that déjà vu and IAMs are related phenomena which are central to understanding a person's relationship with their personal past. Future research may identify how we should think about the role of gist extraction, pattern matching or whether there is one type of phenomenological familiarity or two, but at this moment, we would like our field to converge on the idea of core and attributional systems which are critical for examining our dimensions of content and feelings which clearly make up our experiences of the personal past. Central to this idea is the notion that there is a constantly active system surveying the external environment and assessing its relation to the personal past. This proposes that anything in the external world could possibly act as a cue. The extent to which this cue arouses feelings or retrieval of content from the personal past can be thought of using a threshold mechanism, and the quality and classification of what comes to mind rests on what level of activation the cue affords, and how this activation is interpreted. We believe that, as suggested by Li, Jones, & Laird (Li et al.), in future we will be able to develop computational models of both IAMs and déjà vu that may allow testing at least some ideas presented in our framework. We welcome such a possibility with interest, especially since Li et al. already successfully modelled IAMs and prospective memory.

As a final remark, we are convinced that discussing déjà vu in the context of autobiographical memory research and its theory, especially given the fact that personal past is a key context in which it occurs, may advance our understanding of déjà vu. Over the years, déjà vu was not even close to being linked with autobiographical memory although it clearly is about the personal past. We believe that it is time to deepen our understanding of not only déjà vu but also IAMs. There are plenty of questions to be answered. Some of which, we argue, can only be solved by taking a new perspective based on thinking about memory retrieval being rich and multidimensional, and above all dynamic. In their commentary, Markowitsch, Kordon, & Staniloiu remind us of the classical conceptualisation of memory systems. In our framework, we quite deliberately and extensively discussed the dynamic aspects of memory retrieval processes which are not system-unique. If nothing else, studying déjà vu and involuntary memories should help us shift focus away from this traditional systems approach onto a more nuanced, ever-active, person-centred view of the cognitive neuroscience of memory retrieval.

Financial support. The paper was supported by the National Science Centre, Poland (K. B., grant number UMO-2019/35/B/HS6/00528) and the Institut Universitaire de France (C. M.).

Note

1. Nonetheless, of those people who reported having had the experience, the frequency of déjà vu was related to IAMs (non-parametric correlations, $r[618] = 0.16$, $p < 0.001$ [weekly] and $r[620] = 0.24$, $p < 0.001$ [yearly]).

References

- Ball, C. T., & Little, J. C. (2006). A comparison of involuntary autobiographical memory retrievals. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(9), 1167–1179. <https://doi.org/10.1002/acp.1264>
- Barzykowski, K., & Mazzoni, G. (2022). Do intuitive ideas of the qualities that should characterize involuntary and voluntary memories affect their classification? *Psychological Research*, 86, 170–195. <https://doi.org/10.1007/s00426-020-01465-3>
- Barzykowski, K., & Niedźwieńska, A. (2018). Priming involuntary autobiographical memories in the lab. *Memory*, 26(2), 277–289. <https://doi.org/10.1080/09658211.2017.1353102>
- Barzykowski, K., Niedźwieńska, A., & Mazzoni, G. (2019). How intention to retrieve a memory and expectation that it will happen influence retrieval of autobiographical memories. *Consciousness and Cognition*, 72, 31–48. <https://doi.org/10.1016/j.concog.2019.03.011>
- Barzykowski, K., & Staugaard, S. R. (2018). How intention and monitoring your thoughts influence characteristics of autobiographical memories. *British Journal of Psychology*, 109(2), 321–340. doi:10.1111/bjop.12259
- Bastin, C., Besson, G., Simon, J., Delhay, E., Geurten, M., Willems, S., & Salmon, E. (2019). An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, 42, e281. <http://doi.org/10.1017/S0140525X19000621>
- Brown, A. S. (2003). A review of the déjà vu experience. *Psychological Bulletin*, 129, 394–413.
- Cleary, A. M. (2008). Recognition memory, familiarity, and déjà vu experiences. *Current Directions in Psychological Science*, 17(5), 353–357.
- Curot, J., Pariente, J., Hupé, J. M., Lotterre, J. A., Mirabel, H., & Barbeau, E. J. (2021). Déjà vu and prescience in a case of severe episodic amnesia following bilateral hippocampal lesions. *Memory*, 29(7), 843–858. <https://doi.org/10.1080/09658211.2019.1673426>
- El Haj, M., Colombel, F., Kapogiannis, D., & Gallouj, K. (2020). False memory in Alzheimer's disease. *Behavioural Neurology*, 2020, 5284504. <https://doi.org/10.1155/2020/5284504>
- Feinberg, T. E., & Shapiro, R. M. (1989). Misidentification–reduplication and the right hemisphere. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, 2(1), 39–48.
- Funkhouser, A. (1995). Three types of déjà vu. *Scientific and Medical Network Review*, 57, 20–22.
- Gallo, D. A., Sullivan, A. L., Daffner, K. R., Schacter, D. L., & Budson, A. E. (2004). Associative recognition in Alzheimer's disease: Evidence for impaired recall-to-recognize. *Neuropsychology*, 18(3), 556–563. <https://doi.org/10.1037/0894-4105.18.3.556>
- Kaefer, K., Stella, F., McNaughton, B. L., & Battaglia, F. P. (2022). Replay, the default mode network and the cascaded memory systems model. *Nature Reviews Neuroscience*, 1–13.
- Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, 207(4430), 557–558. <https://doi.org/10.1126/science.7352271>
- Kvavilashvili, L., Niedźwieńska, A., Gilbert, S. J., & Markostamou, I. (2020). Deficits in spontaneous cognition as an early marker of Alzheimer's disease. *Trends in Cognitive Sciences*, 24(4), 285–301.
- Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind & Language*, 15(1), 184–218. <https://doi.org/10.1111/1468-0017.00129>
- Leeds, M. (1944). One form of paramnesia: The illusion of déjà vu. *Journal of the American Society for Psychical Research*, 38, 24–42.
- Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract cuing: The Proustian view is incorrect. *Applied Cognitive Psychology*, 18(7), 893–899. <https://doi.org/10.1002/acp.1020>
- Mace, J. H. (2005). Priming involuntary autobiographical memories. *Memory*, 13, 874–884.
- Maillet, D., & Schacter, D. L. (2016). From mind wandering to involuntary retrieval: Age-related differences in spontaneous cognitive processes. *Neuropsychologia*, 80, 142–156. <https://doi.org/10.1016/j.neuropsychologia.2015.11.017>
- Montaldi, D., & Kafkas, A. (2022, September 7). Neural mechanisms of familiarity. <https://doi.org/10.31234/osf.io/zfq5s>
- Moulin, C. J. (2013). Disordered recognition memory: Recollective confabulation. *Cortex*, 49(6), 1541–1552.
- Moulin, C. J. A. (2018). *The cognitive neuropsychology of déjà vu. Essays in cognitive psychology*. Psychology Press (Routledge).
- Moulin, C. J. A., Carreras, F., & Barzykowski, K. (2023). The phenomenology of autobiographical retrieval. *Wiley Interdisciplinary Reviews: Cognitive Science*, 14(3), e1638. <https://doi.org/10.1002/wcs.1638>
- Moulin, C. J. A., Conway, M. A., Thompson, R. G., James, N., & Jones, R. W. (2005). Disordered memory awareness: Recollective confabulation in two cases of persistent déjà vu. *Neuropsychologia*, 43(9), 1362–1378. <https://doi.org/10.1016/j.neuropsychologia.2004.12.008>
- Moulin, C. J. A., Souchay, C., Bradley, R., Buchanan, S., Karadoller, D. Z., & Akan, M. (2014). Déjà vu in older adults. In B. Schwartz & A. Brown (Eds.), *The tip of the tongue states and related phenomena* (pp. 281–304). Cambridge University Press.
- Moulin, C. J. A., Bell, N., Turunen, M., Baharin, A., & O'Connor, A. R. (2021). The induction of jamais vu in the laboratory: Word alienation and semantic satiation. *Memory*, 29(7), 933–942. doi:10.1080/09658211.2020.1727519
- Norman, D. A., & Bobrow, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, 11(1), 107–123. [https://doi.org/10.1016/0010-0285\(79\)90006-9](https://doi.org/10.1016/0010-0285(79)90006-9)
- Parker, E. S., Cahill, L., & McGaugh, J. L. (2006). A case of unusual autobiographical remembering. *Neurocase*, 12(1), 35–49. <https://doi.org/10.1080/13554790500473680>
- Perrin, D., Moulin, C. J. A., & Sant'Anna, A. (2023). Déjà vécu is not déjà vu: An ability view. *Philosophical Psychology*, 1–31. <http://doi.org/10.1080/09515089.2022.2161357>
- Ryom, K. I., Stendardi, D., Ciaramelli, E., & Treves, A. (2022). Computational constraints on the associative recall of spatial scenes. *bioRxiv*.
- Souchay, C. (2007). Metamemory in Alzheimer's disease. *Cortex*, 43(7), 987–1003. [https://doi.org/10.1016/s0010-9452\(08\)70696-8](https://doi.org/10.1016/s0010-9452(08)70696-8)

- Souchay, C., Isingrini, M., & Gil, R. (2002). Alzheimer's disease and feeling-of-knowing in episodic memory. *Neuropsychologia*, *40*(13), 2386–2396. [https://doi.org/10.1016/S0028-3932\(02\)00075-1](https://doi.org/10.1016/S0028-3932(02)00075-1)
- Souchay, C., & Moulin, C. J. (2009). Memory and consciousness in Alzheimer's disease. *Current Alzheimer Research*, *6*(3), 186–195. <https://doi.org/10.2174/156720509788486545>
- Urquhart, J. A., Sivakumaran, M. H., Macfarlane, J. A., & O'Connor, A. R. (2021). fMRI evidence supporting the role of memory conflict in the déjà vu experience. *Memory*, *29*(7), 921–932. <https://doi.org/10.1080/09658211.2018.1524496>
- Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy- attribution account of feelings of familiarity. *Acta Psychologica*, *98*, 141–165.
- Yeung, R. C., & Fernandes, M. A. (2021). Recurrent involuntary memories are modulated by age and linked to mental health. *Psychology and Aging*, *36*(7), 883.
- Zareen, G., Ahraf, F., Barzykowski, K., & Moulin, C. J. A. (in preparation). Descriptions and incidence of déjà vu and other spontaneous cognitive phenomena in Pakistan, Poland and France.