

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., Delhaye, 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/https://doi.org/10.1016/j.cortex.2022.02.008>
- Chua, E., Pergolizzi, D., & Weintraub, R. R. (2014). The cognitive neuroscience of metamemory 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/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., Pernia-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., & Fenn, 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?FirstName=Anne&LastName=Cleary>
<https://www.cehd.umn.edu/dpsych/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 experiencer.

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 experiencer). 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