Please cite as: Veit, W. & Browning, H. (2022). Evolutionary Mismatch and Anomalies in the Memory System. Preprint.

Check www.walterveit.com for citation details once published

Evolutionary Mismatch and Anomalies in the Memory System

Walter Veit

https://orcid.org/0000-0001-7701-8995

University of Bristol

Department of Philosophy, University of Bristol, BS6 6JL UK

wrwveit@gmail.com

https://walterveit.com/

Heather Browning

https://orcid.org/0000-0003-1554-7052

University of Southampton

Department of Philosophy, University of Southampton, UK

DrHeatherBrowning@gmail.com

https://www.heatherbrowning.net/

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.

Main text:

Barzykowski and Moulin (2022) 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. Barzykowski and Moulin (2022) 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 Stearns 2012; Manus 2018; 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 doesn't 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 favor 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 Barzykowski and Moulin (2022) 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 forthcoming). 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 Barzykowski and Moulin (2022). 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 skeptical 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.

Conflict of interest statement:

The authors have no conflicts of interest to report.

Funding Information:

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).

References

Barzykowski, K. & Moulin, C. (2022). Are Involuntary Autobiographical Memory and Déjà Vu Natural Products of Memory Retrieval? Behavioral and Brain Sciences.

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. https://doi.org/10.1080/10888705.2021.2014838

Veit, W. and 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. (forthcoming). Hominin Life History, Pathological Complexity, and the Evolution of Anxiety. Behavioral and Brain Sciences.