



„Analiza i Egzystencja” 65 (2024), 21-34
ISSN (print): 1734-9923
ISSN (online): 2300-7621
DOI: 10.18276/aie.2024.65-02

ARTYKUŁY

LUIZ DO VALLE MIRANDA

ORCID:0000-0003-1838-5693

Charles University in Prague

email: luizdovalle2@gmail.com

Memory Traces, Phenomenology and the Simulationist vs Causal Theory Dispute

Keywords: memory traces, familiarity, causal connection

Słowa kluczowe: ślady pamięciowe, swojskość, związek przyczynowy

Abstract

Philosophy of memory is a hot topic in the cognitive sciences and philosophy of mind. This work examines the dispute between simulations and causal theories of memory by means of an examination of the feeling of familiarity and its relation to memory traces, more specifically how they increase the fluency of reconstruction of past episodes. Understanding the relationship between familiarity and memory traces and, furthermore, the relation between a fully-fledged phenomenology of memory and a sense of subjective certainty of the episode has occurred in the past, lead to a different interpretation of the rivalry between the CTM and the simulationist account.

Introduction

Our tendency is to believe that everything we remember has actually been experienced by us and that our remembering of this event, through a faculty that came to be known as Episodic Memory (Tulving, 1972), is possible given the fact that we have experienced this episode. Whenever we remember

an episode that, in fact, has not occurred in our personal past, it is intuitive to call it an error—confabulation, misremembering, etc.

In order to address to what extent these errors should be considered a failure of the memory system, how they are produced, and how they relate to the nature of memory in general, philosophers have increasingly concerned themselves with understanding this pervasive cognitive capacity. Philosophy of memory, therefore, has become one of the hot topics in the cognitive sciences and the philosophy of mind (Bernecker & Michaelian, 2017).

Most of the contemporary discussion on the philosophy of memory has as its starting point Martin and Deutscher's (1966) theory of memory. This essay is no different. Their Causal Theory of Memory (CTM) has for decades been considered as the canon for thinking and discussing the nature of memory. However, recent empirical research¹ on the constructive character of memory has served as the basis for challenges to the CTM, and given space for the proposal of new approaches toward studying and defining such capacity (Michaelian, 2011).

One of the most prominent theories of memory proposed in recent years is Michaelian's (2016) simulationist theory, which completely challenges the CTM's main tenet that remembering a past episode requires a causal connection between that event and the act of remembering and that this connection is sustained by a memory trace (Robins, 2016). For the simulationist, given the constructive nature of memory, i.e., the fact that the content stored between an episode and the act of remembering can change to varying degrees, remembering an episode does not necessarily involve a causal connection to this episode and, furthermore, memory and imagination are

¹ The connection between empirical research and philosophical theorizing is *per se* problematic. Hereon, all theories and problems are to be taken as philosophical in nature, problems to be discussed and reflected with the goal of presenting an interpretation of the relation between the entities under discussion that can be taken up for further deliberation, or perhaps for informing scientific practice. Results from empirical research are important starting points for the formulation of arguments in favor of a certain position; however, they are one of many different illustrative methods to strengthen a philosophical argument. In this sense, the present approach to philosophical research aligns to Chalmers', who asserts that even though science is a terrific way for expanding one's imagination while doing philosophy, enduring questions cannot be definitively resolved by relying solely on empirical findings (Grim, 2009). For further discussion on the issues related to an empirically informed philosophy of mind, see Irvine (2014).

not to be strictly distinguished as they would be two sides of the same capacity—an episode construction system.

With these two ways of conceiving episodic memory in mind, this essay sets out to present a critique of Michaelian's simulationist theory. In this essay, I advance the position that the CTM and the simulationist theory are not mutually exclusive theories of memory and that they can be understood as being part of a spectrum of memory phenomena in relation to involving memory traces during the act of remembering (recall).

Furthermore, this difference between the involvement of memory trace is reflected in the phenomenology of memory, more specifically in the feeling of familiarity. I argue that Michaelian's theory cannot fully account for the phenomenological aspects, as it leaves familiarity out of the picture and conflates the meaning of auto-noesis. Nevertheless, it looks like his version of an episodic construction system sketches important aspects of memory. Considering the activity of such a system to produce representations of the past to be an instance of episodic memory thus depends on which aspects one holds fundamental in defining the latter capacity.

From the Causal to the Simulation Theory

As mentioned in the introduction, the investigation begins with an exploration of the classical causal theory of memory, then explore some findings about the constructive aspect of memory and how they can be accommodated by a causal theory, culminating in the formulation of the simulationist theory. The exploration is based on the content of Kourken Michaelian's *Mental Time Travel*.

The causal theory of memory, as originally formulated by Martin and Deutscher (1966), is based on the presupposition that for remembering there must necessarily exist a causal connection between a subject's previous representation R' and the current representation R brought about by the act of remembering. However, to exclude some cases of forgetting and thereafter reacquiring such representation, more must be added to the theory in relation to this causal connection.

In order to rule out such cases, the notion of a memory trace is introduced. Despite the divergent ways of conceiving memory traces (for a further discussion of this topic, see e.g., Robins, 2016 and de Brigard, 2014),

in the classical formulation of the CTM it should be understood as an exact copy of R' . Furthermore, another necessary condition for remembering, according to the CTM, is that the memory trace has been continuously stored between the episode and the time of recall.

The condition that the memory trace must be an unaltered copy of the initial representation, however, is intrinsically problematic, as minor transformations naturally occur between R' and R that we do not rule out as cases of remembering, such as generalizations and deletions of aspects of the representations. An example of research on constructive memory is that of Alba and Hasher (1983), who suggest four possible dimensions in which memory can be constructive: *selection* of details at the encoding of R' , *abstraction* from certain features, *interpretation* given prior knowledge, and *integration* of the former modifications in a coherent whole. Moreover, they suggest that during the retrieval of R , a process of reconstruction can be at play in which information from other sources, such as semantic memory or contextual information, can be included in the representation.

Employing such empirical illustrations of more extensive possibilities for variations in the memory trace, CTM, as a philosophy of memory, departs from the classical version into a constructive version that accepts that the memory trace and the representation R must not be exactly the same as R' but can vary—increase, decrease or even merge the content of different experiences—to a certain extent. The extrapolation of how much R can differ from R' and therefore the need and the place of memory traces if this divergence can be extrapolated leads to the possibility and the suggestion of the simulation theory of memory.

Drawing on Tulving's (2001) developments on the definition of episodic memory towards interpreting it as a function of a “mental time travel” system responsible for both re-experiencing past episodes, as well as “pre-experiencing” or simulating future episodes, Michaelian starts proposing his own version of episodic memory as one part of a more general episodic construction system.

According to Michaelian, we possess an episodic construction system that has as its proper function drawing on acquired information to construct or simulate novel representations, either of past or future episodes. This system, therefore, is essentially constructive, drawing not only on experiential information to construct such representations but also on non-experiential information possessed by the subject, for example, semantic memory.

The schism between the simulation theory and the constructive CTM lies in their treatment of the necessity of memory traces, as the former does not share the assumption that at least some information originating in the experience of the remembered episode must be used in constructing the representation of the episode. Summing up, according to the simulation theory we are endowed with an episodic construction system that has different functions, including simulating the past. Remembering, thus, can be defined as when a representation R:

is produced by a properly functioning episodic construction system which aims to produce a representation of an episode belonging to [a subject] S's personal past. (Michaelian, 2016, p. 105)

Taking into consideration Michaelian's analysis of the conditions for remembering that are postulated by the CTM and the simulationist theory, it looks like these theories are markedly different and completely incompatible. On the one hand, according to Michaelian's reconstruction of the CTM, remembering depends on acquiring some representation R', maintaining it continuously stored in the form of a memory trace that serves as a copy of it (allowing minor transformations), and retrieving a representation R through this trace that must be similar to R' to a certain extent. On the other hand, for the simulationist theory, remembering only consists of constructing a representation R of a past episode through a functioning episodic construction system.

Presented in this way, they do indeed look like two opposing theories. However, if the presentation changes from the conditions for remembering to the dispositions involved in remembering, the differences between these theories start to diminish. Reconstructing the classical CTM in terms of dispositions to remember, it can be understood that remembering only happens when recall prompts (voluntarily or involuntarily) activate a system responsible for retrieving the representation R of a past episode exclusively through the dispositional character of memory traces.

Furthermore, according to the constructivist CTM, remembering happens when recall prompts activate a system aiming to reconstruct a representation R drawing on contextual and stored information—necessarily including information stored via memory traces. Finally, for the simulationist theory, the dispositions that lead the episode construction system to produce

a representation R depend on stored information, as well as contextual information, not necessarily needing memory traces of past episodic experiences.

Accepting the fact that representations of past episodes are produced by an episodic construction system and that memory traces might not even be involved in such a construction, however, does not mean that talking about memory traces and both the classic and constructivist CTM should be dropped. Moreover, the mere possibility of producing a representation R with the content of a past episode does not entail that this is a case of episodic memory, as the act of remembering does not involve only the generation of content, but also a range of phenomenological qualities. More must be said about the phenomenology of memory and how it is related to this episodic construction system and, in the end, to memory traces in general.

Memory Traces, Auto-noesis and Familiarity

There is a characteristic phenomenology of remembering episodes, i.e., it feels like something to be engaged with episodic memory. Tulving (1985) was the first to divide different memory systems according to their phenomenology. Procedural memory was supposedly *anoetic*, i.e., its performance did not require consciousness. Semantic memory was considered *noetic*, i.e., it is brought to consciousness, but the circumstances of the acquisition of such knowledge are lost. Episodic memory is supposed to be *auto-noetic*, which means that the episode is brought into consciousness together with information about its original experience.

Invoking the concept of auto-noesis to explain the phenomenology of episodic memory is controversial, as the very definition and use of such a term have been widely divergent among authors. On the one hand, Klein (2015) focuses on auto-noetic awareness as some kind of feeling of warmth and intimacy, as illustrated by this passage, “[The patient R.B.] lacked the warmth, intimacy, and feeling of reliving associated with auto-noetic experience” (p. 18), while Michaelian (2016) highlights auto-noesis as “[enabling] the agent to discriminate between self- and other-oriented forms of episodic imagination” (p. 232). On the other hand, Tulving’s (1993) original formulation focused on auto-noesis as “a unique awareness of re-experiencing here and now something that happened before, at another time and in another place” (p. 68).

Generalizing auto-noesis does not do justice to a more nuanced investigation of the phenomenology of memory. A more thorough investigation of the experiential states associated with the act of remembering must be performed, and that indeed is attempted by Michaelian, who divides the phenomenology of memory into a feeling of pastness and a feeling of self-presence (what he calls auto-noesis).

The first feeling indeed can differentiate between how past and future-oriented forms of episodic construction are presented to consciousness, while the second discriminates how episodic construction from the subject's self-point of view can be different from others-based episodes. However, as Michaelian himself admits, these inner feelings alone cannot account for the different ways that the construction of episodes belonging to the personal past and the counterfactual past are presented and, therefore, do not exhaust the distinctiveness of the phenomenology of episodic memory.

Further philosophical research on identifying the distinctive phenomenology involved in remembering particular events as past experiences suggests that there is another feeling characteristic to such capacity, i.e., a feeling of familiarity (Teroni, 2017). Even though a thorough definition of such a quality is not yet sketched, it can be understood as a feeling of fluency in the construction of a representation stemming from the fact that it has been already encountered in past experiences.

Michaelian (2016) lays out some concerns on the usefulness of talking about familiarity. When distinguishing the possible phenomenology that could be the source of monitoring for the subject's awareness of which function the episodic construction system is performing, Michaelian rejects the idea that familiarity might play a role in bringing about a feeling of pastness and therefore distinguishing memory and imagination. He concludes that familiarity may have limited validity in distinguishing remembering from other forms of construction.

However, I consider Michaelian's dismissal of the usefulness and of a deeper discussion of familiarity to be too forward and precocious. Given the pervasiveness of this feeling when we indeed remember an episode, more attention must be given to it as a distinctive aspect of remembering episodes, as discussed by Teroni (2017), especially when working together with the two feelings mentioned above: pastness and self-presence.

As discussed by Michaelian, familiarity is neither necessary nor sufficient for pastness; however, it serves as a cue that the constructed episode has been

previously experienced, so that it is a product of remembering. This familiarity can be related to the fluency of construction, as remembering should be less effortful than other forms of construction, given that the produced representations are more similar to the stored information.

As discussed above, episodic memory and episodic counterfactual thinking cannot be distinguished solely by differences in phenomenology if only pastness and self-presence are considered. The feeling of fluency of construction related to familiarity can serve to distinguish these two capacities, as the construction of episodic counterfactual representation is indeed more effortful as it cannot be helped by any kind of similarly stored representation—i.e., memory traces—while episodic memory possesses this kind of familiarity, that together with the feeling of pastness and self-presence sum up to the feeling of “warmth and intimacy” long presented by authors such as Russell (1921) and James (1890) as the distinctive character of episodic memory.

The fact that memory traces and familiarity are intertwined is not something novel. Tulving (1985) proposed evidence for a relation between the strength of memory traces and his conception of auto-noetic consciousness, which also involves a sense of familiarity. Furthermore, the very definition of the feeling of familiarity as a measure of the fluency of recollection favors a positive correlation between the presence and strength of memory traces and the intensity of this feeling.

Considering two cases of episodic reconstruction, case 1 where there are a number n of dispositions to bring about the reconstruction of an episode into a representation, and case 2 where there are the same n dispositions plus the dispositions that form the memory trace, it is obvious that in the second case, the fluency of recollection will be higher. Not only that but considering memory traces as structural analogs of the original representation—to some extent allowing transformations of them—even if the dispositions n in case 2 are to a certain degree weaker, the fluency of recollection is still higher than in case 1, given the fact that the representation to be reconstructed is already partially present as a memory trace.

Evidence for this can be found in one’s own reflection of familiarity given a range of cases of remembering. Comparing an important and detailed event seems more familiar when remembered than a temporally more distant episode. Moreover, considering the case of everyday activities, such as walking home from work, the act of remembering them involves multiple

merged memory traces from every time such an event happens and hence has the most feeling of familiarity.

Going back to Michaelian's framework, it is possible to sketch the importance of familiarity to episodic memory, and therefore the consequences of the lack of it. He draws two roles for phenomenology in his account of the episodic construction system. The first one, as already discussed, is to distinguish what kind of episodic construction the subject is involved in, whether it is imagination, past counterfactual thinking, episodic memory, etc. They feel different, and these different feelings serve as indicators for the subject that he is engaged in one, and not in the other type of thinking. However, without including familiarity as one of these feelings, the problem of distinguishing between episodic memory and episodic counterfactual thinking solely on the basis of feelings of pastness and self-presence appears.

Not only does the experience of remembering play a role in the episodic construction system in general, but it has an important place in episodic memory as such. Drawing on Klein (2013) and Tulving (1985), Michaelian summarizes their ideas in the following passage:

Autonoesis, on this view, functions as a means of reducing uncertainty about whether or not an apparently remembered event actually occurred: by providing the agent with a sense that the remembered event belongs to his personal past, it provides him with a reason for accepting, and acting on, the retrieved content. (p. 228)

Even though Michaelian considers the truth of Klein's and Tulving's approach to autonoesis, Michaelian rejects it on the basis that autonoesis alone cannot account for this subjective certainty and argues that it is the full phenomenology that plays a role in bringing about this certainty. Nevertheless, he is not able to sketch a full picture of how phenomenology alone brings about subjective certainty and appeals that the content of the constructed representation also plays a complementary role.

Defining memory, though, on the basis of the content of the remembered representation is problematic, as argued by Klein (2015). Memory should be understood on the basis of the manner in which this content is experienced. Michaelian, however, is unable to give a satisfying account of the subjective certainty characteristic of memory in terms of the way the content presents itself to awareness.

Klein's (2015) discussion on memory provides one way to evaluate whether or not the CTM, Michaelian's simulationist theory, or any kind of theory of memory involving an episodic construction system really is able to tackle the problem of memory presented in this article. Taking as a starting point the existence of an episodic construction system, and from there evaluating the role and necessity of memory traces, it is possible to give a conclusive assessment of Michaelian's simulation theory.

The episodic construction system presented by Michaelian is capable of producing representations of episodes in a range of different ways, and the constructed representations are accompanied by a different phenomenality depending if they are instances of imagination, episodic memory, episodic counterfactual thinking, etc. The construction of such representations is brought about by the interaction of cue triggers from the environment and internal dispositions to produce these representations.

Among these dispositions, some are formed from non-experiential information, such as semantic memory, but also from experiential information. Pieces of experiential information might or might not be memory traces—stored representations of an episode that can be slightly transformed but not to the point of rendering the initial episode unrecognizable. Remembering an episode, according to Michaelian, consists in the production of a representation of a past episode by a properly functioning episodic construction system. The dispositions involved in producing such a representation can, but do not necessarily, involve memory traces. However, such a way of understanding episodic memory is problematic as Michaelian's account does not take into consideration an important part of "what it feels like" to remember an episode, i.e., the sense of familiarity. The absence of familiarity is connected to the dispensability of the inclusion of memory traces in his philosophy of memory.

This disregard for familiarity, thus, has some consequences in his account of a part of episodic memory that is considered by some authors as of extreme importance, i.e., a fully-fledged subjective certainty that the remembered event indeed has occurred in a person's past, giving this person full acceptance on the content of such representation and enough reason to act according to it. Even though some sense of certainty can be produced together with the feelings of self-presence and pastness discussed by Michaelian, they are of a much lesser degree than what is described to accompany the sense of familiarity.

The notion of degree in discussing the difference between episodes reconstructed on the basis of dispositions involving memory traces and dispositions not involving them is the last theme of this paper. It looks like the classical CTM, constructivist CTM and simulationist theory are not different theories in kind, but only in degree. Each theory elucidates a different aspect of episodic memory, and to consider the case elucidated by each one as paradigmatic of remembering depends on which goals and which aspects of episodic remembering an author has in mind as the most important.

While accepting some kind of classical CTM rules out instances of reconstruction of episodes as cases of memory that the simulationist may account for, the adherents of the CTM maintain subjective certainty and epistemic authority as well as a fully-fledged phenomenology of the retrieved representations. On the other hand, the description of representations that the simulationist may accept as cases of episodic memory may not necessarily be accompanied by such a subjective experience, but surely many more cases of remembering are incorporated into this theory.

Lastly, I am not claiming some kind of relativism towards theories of memory, and that no theory will ever be able to account for the whole nuances of episodic memory and memory experience, just that adhering to either the CTM or simulationism depends on what aspect of remembering one hold to be more important. In the case of the CTM, this aspect is the subjective certainty and epistemic authority accompanied by episodic memory, while for the simulationist theory, it is the inclusion of a larger number of cases of reconstruction of past episodes.

Conclusion

A closer look at the phenomenological discussion presented in Michaelian's *Mental Time Travel* shows a conflation between auto-noetic awareness and the feeling of self-memory. Further investigation into this confusion elucidates a pervading theme in the philosophy of memory, i.e., a lack of agreement on the notion of auto-noesis and therefore a range of different usages for such a term.

From this starting point, the solution for investigating the phenomenology of memory has been to divide the different ways in which authors refer

to such terms as different aspects of the phenomenology of memory, which were considered to be a feeling of pastness, self-presence and familiarity. While Michaelian's simulationist theory accounts for the first two, it dismisses and fails to include familiarity in its phenomenological picture.

A further examination of the feeling of familiarity indicates a close relation to memory traces, as the presence of such an experience is accompanied by fluency in the reconstruction of past episodes. This investigation leads the way to understanding the controversies between a CTM and the simulationist theory, as their main difference is in how they consider the involvement of memory traces.

Understanding the relationship between familiarity and memory traces and, furthermore, the relation between a fully-fledged phenomenology of memory and a sense of subjective certainty of past episodes leads to a different interpretation of the rivalry between the CTM and the simulationist account. It looks like they are elucidating different aspects of memory, and that, therefore, one's adherence to each of the theories is connected to what they hold as canonical and what they want to explain with such a conception of episodic memory.

Finally, this is not to say that a theory that encompasses all the nuances of episodic memory is unachievable, but more as a warning not to treat the causal theory and the simulationist theory as two antagonists in the study of memory. Hopefully, this warning is able to influence how philosophy of memory treats cases of successfully remembering and how it differs from memory errors.

References

- Alba, J. W., Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93 (2), 203–231. <https://doi.org/10.1037/0033-2909.93.2.203>.
- De Brigard, F. (2014). The nature of memory traces. *Philosophy Compass*, 9 (6), 402. <https://doi.org/10.1111/phc3.12133>.
- Grim, P. (2009). *Mind and consciousness: 5 questions*. Automatic Press.
- Irvine, E. (2014). Problems and possibilities for empirically informed philosophy of mind. In: M. Sprevak, J. Kallestrup (Eds.), *New waves in philosophy of mind: New waves in philosophy*. Palgrave Macmillan. http://dx.doi.org/10.1057/9781137286734_9.

- James, W. (1890). *The principles of psychology*. Macmillan.
- Klein, S. B. (2013). The complex act of projecting oneself into the future. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4 (1), 63–79. <https://doi.org/10.1002/wcs.1210>.
- Klein, S. B. (2015). What memory is. *WIREs: Cognitive Science*, 6 (1), 1. <https://doi.org/10.1002/wcs.1333>.
- Mahr, J. B., Csibra, G. (2018). Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences*, 41, 1–93. <https://doi.org/10.1017/S0140525X17000012>.
- Martin, C. B., Deutscher, M. (1966). Remembering. *Philosophical Review*, 75, 161–196. <https://doi.org/10.2307/2183082>.
- Michaelian, K. (2011). Generative memory. *Philosophical Psychology*, 24 (3), 323–342. <https://doi.org/10.1080/09515089.2011.559623>.
- Michaelian, K. (2016). *Mental time travel: episodic memory and our knowledge of the personal past*. The MIT Press. <https://doi.org/10.7551/mitpress/10591.001.0001>.
- Michaelian, K., Bernecker, S. (2017). *The Routledge handbook of philosophy of memory*. Routledge, Taylor & Francis Group. <https://doi.org/10.4324/9781315687315>.
- Robins, S. (2016). Representing the past: memory traces and the causal theory of memory. *Philosophical Studies*, 173 (11), 2993–3013. <https://doi.org/10.1007/s11098-016-0647-x>.
- Russell, B. (1921). *The analysis of mind*. George Allen & Unwin.
- Teroni, F. (2017). The phenomenology of memory. In: S. Bernecker, K. Michaelian (Eds.), *Oxford handbook of philosophy of memory* (pp. 21–33). Oxford University Press.
- Tulving, E. (1972). Episodic and semantic memory. In: E. Tulving, W. Donaldson (Eds.), *Organization of memory* (pp. 381–402). Academic Press.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26 (1), 1–12. <https://doi.org/10.1037/h0080017>.
- Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, 2 (3), 67–70.
- Tulving, E. (2001). Origin of autoevidence in episodic memory. In: H. L. Roediger III, J. S. Nairne, I. Neath, A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 17–34). American Psychological Association.

Author Note

Luiz do Valle Miranda is a Ph.D. candidate in philosophy of education at Charles University in Prague. He is interested in topics ranging from philosophy of mind and psychology to education and technology. Besides these fields, he has published works in the philosophy of childhood and evolution. He has also spoken at several conferences on educational research.

Address for correspondence: luizdovalle2@gmail.com

Citation

do Valle Miranda, L. Memory traces, phenomenology and the simulationist vs causal theory dispute. *Analiza i Egzystencja*, 65 (1), 21–34. <https://doi.org/10.18276/aie.2024.65-02>.