



Causation and mnemonic roles: on Fernández's Functionalism

Nikola Andonovski

► To cite this version:

Nikola Andonovski. Causation and mnemonic roles: on Fernández's Functionalism. *Estudios de Filosofía Practica e Historia de las Ideas*, 2021, 64, pp.139-153. <10.17533/udea.ef.n64a07>. <hal-04858778>

HAL Id: hal-04858778

<https://hal.science/hal-04858778v1>

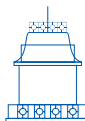
Submitted on 30 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



ARTÍCULO
DE REFLEXIÓN

Causation and mnemonic roles: on Fernández's Functionalism*

Nikola Andonovski

Université Grenoble Alpes, Grenoble, France

E-mail: ni.andonovski@gmail.com

Recibido: 15 de enero de 2021 | Aceptado: 25 de mayo de 2021

<https://doi.org/10.17533/udea.ef.n64a07>

Abstract: Debates about causation have dominated recent philosophy of memory. While causal theorists have argued that an appropriate causal connection to a past experience is necessary for remembering, their opponents have argued that this necessity condition needs to be relaxed. Recently, Jordi Fernández (2018; 2019) has attempted to provide such a relaxation. On his functionalist theory of remembering, a given state need not be caused by a past experience to qualify as a memory; it only has to realize the relevant functional role in the subject's mental economy. In this comment, I argue that Fernández's theory doesn't advance the debate about memory causation. I propose that this debate is best understood as being about the existence of systems, which support kinds of interactions that map onto the relations dictated by (causal) theories. Since Fernández's functionalism tells us very little about this empirical question, the theoretical gains from endorsing it are minimal.

Keywords: episodic memory, functionalism, mnemonic role, realization, *ceteris paribus* laws

* This paper is part of a project investigating causation in memory, conducted at the Centre for Philosophy of Memory at the Université Grenoble Alpes.

Cómo citar este artículo

Andonovski, N. (2021). Causation and mnemonic roles: on Fernández's Functionalism. *Estudios de Filosofía*, 64, 139-153. <https://doi.org/10.17533/udea.ef.n64a07>



Causalidad y roles mnemónicos: sobre el funcionalismo de Fernández

Resumen: Los debates sobre la causalidad han dominado la reciente filosofía de la memoria. Mientras que los teóricos causales han argumentado que para recordar es necesaria una conexión causal apropiada con una experiencia pasada, sus oponentes han argumentado que esta condición necesita ser relajada. Recientemente, Jordi Fernández (2018; 2019) ha intentado esto. Según su teoría funcionalista del recuerdo, un estado dado no necesita ser causado por una experiencia pasada para calificar como un recuerdo; sólo tiene que realizar el papel funcional relevante en la economía mental del sujeto. En este comentario, sostengo que la teoría de Fernández no avanza en el debate sobre la causalidad de la memoria. Propongo que este debate se entiende mejor como si tratara acerca de la existencia de sistemas, que realizan los tipos de interacciones que corresponden a las relaciones dictadas por las teorías (causales). Dado que el funcionalismo de Fernández nos dice muy poco sobre esta cuestión empírica, los beneficios teóricos de respaldarla son mínimos.

Palabras clave: memoria episódica, funcionalismo, rol mnémico, realización, leyes ceteris paribus

Nikola Andonovski

He received his PhD from the Department of Philosophy at Johns Hopkins University in 2020. He is currently affiliated with the Centre for Philosophy of Memory at the Université Grenoble Alpes. His research lies at the intersection of philosophy and the sciences of memory. He is interested in the representational structure of episodic memory, the nature of consolidation and the relationship between memory and reasoning.

ORCID: 0000-0001-6957-0261



1 Fernández's Functionalism

Debates about causation have dominated analytic philosophy of memory. Causal theorists –old (Martin & Deutscher, 1966) and new (Robins, 2016; 2020a)– have argued that an appropriate causal connection to a past experience is necessary for remembering an event.¹ Their simulationist foes, emboldened by some surprising developments in the sciences of memory, have challenged this claim, insisting that such a causal connection is *not* necessary (Michaelian, 2016a; Michaelian & Sant'Anna, 2019).² While, from 10,000 feet, the two camps seem clearly distinct and well-fortified, a closer look reveals a surprisingly intricate terrain, crisscrossed by a number of unexplored routes between them (Andonovski, 2021; Langland-Hassan, 2021). Still, there is a growing suspicion that, in light of the empirical developments –and perhaps also the evolving nature of *theorizing* about memory– a “relaxation of the requirement for a strictly necessary and sufficient condition is welcome” (Lewis, 1966, p. 22).

Recently, Jordi Fernández (2018; 2019) has attempted to provide precisely such a relaxation. Taking inspiration from classic functionalism (Block, 1978; Lewis, 1966), Fernández offers an account on which a mental state qualifies as a memory of a past event just in case it plays the functional role memories play in subjects' cognitive economies. Two kinds of causal relations, according to Fernández, are constitutive of this role: memories tend to cause beliefs about the occurrence of represented events *and* they tend to be caused by past experiences of them. On his Functionalist Theory of Memory (FTM):

[F]or any subject S and event e, S remembers e just in case S [is in a mental state that] tends to cause in S a disposition to believe both that e happened and that S experienced e to happen, and [that] tends to be caused in S by having experienced e to happen (2018, p. 64).^{3,4}

1 Contemporary debaters have dealt, almost exclusively, with *episodic* remembering, which they have characterized as a state/process of entertaining quasi-sensory representations of particular past events. Please keep this in mind while reading the essay. For problems with this dominant view, see Andonovski (2020).

2 In fact, Michaelian (2016a, pp.110-113) denies the necessity of *any* kind of causal connection between memories and represented events, appropriate or otherwise. See also the characterization in Michaelian & Robins (2018).

3 In specifying the functional role of memories, Fernández characterizes them as kinds of mental “images” –i.e. mental states (experiences) that have *phenomenal* properties (see 2019, Ch.1 & 2). I have eschewed talk of “images” for two reasons. First, the phenomenal properties of memories do not matter for my purposes. Second, it seems to me that Fernández shifts somewhat uncomfortably between talking about the mnemonic role (defined in functional/causal terms) and talking about the realizer states, which (necessarily) have phenomenal properties (see note 5). I hope I am not being unfair to him.

4 In his 2019, Fernández characterizes episodic remembering in propositional terms. Accordingly, he talks about remembering *facts* about past experiences/events (pp.47-56). As far as I can see, leaving out this commitment doesn't affect my argument in this essay.

FTM, thus, aims to specify the relevant features in virtue of which a mental state qualifies as a state of remembering. It does so iff it *realizes* the mnemonic role characterized above.⁵ Yet, the theory doesn't aim to exhaustively characterize all features of such realizer states. As Fernández points out, a mental state may be representationally and phenomenologically indistinguishable from a realizer state but fail to play the relevant mnemonic role. Such a state is *not* a state of remembering.

What matters for our purposes, however, is the weakening of the necessity condition favored by causalists. FTM endeavors to do justice to the spirit of causal theories –memories *are* typically caused by past experiences– while seemingly accommodating the possibility of exceptions. Hence, a given state need not be caused by a past experience in order to qualify as a memory; it only has to realize the relevant functional role in the subject's mind. In other words, it has to be the kind of state that *tends to* be caused by past experiences, even if it is not *actually* caused by such an experience. As long as a state “plays the mnemonic role in me, I qualify as remembering the event” (Fernández, 2018, p. 65).⁶ The lesson is a classic one: if we just allow for a little wiggle room and let in occasional exceptions to putative psychological laws, we can accumulate all sorts of theoretical gains.⁷

In this brief comment, I argue that, in the context of memory causation, FTM's theoretical gains are only apparent. To illustrate this, in [section 2](#), I discuss the conditions in which a token mental state counts as a realization of a functional kind. I propose that it does so just in case it is embedded in a system of the right sort. In [section 3](#), I argue that the debate between causalists and simulationists concerns precisely the existence of such systems. Since FTM tells us very little about this, by-and-large empirical, question, it doesn't advance the debate about memory causation. I end the paper by exploring the relation between FTM, causal and simulation theories.

5 Fernández opts for a *role* version of functionalism, according to which remembering is identified with the *second-order* functionally defined property that can be realized by some-or-other realizer state. As Rupert (2006) points out, the literature has not been kind to role functionalism, uncovering a variety of problems and objections. Not the least of these is the so-called “causal exclusion problem”: with the realizer states doing all the relevant causal work, there seems to be no causal work left to be done by the second-order state (see Kim, 1993). Unfortunately, Fernández's account inherits this problem: are memories *epiphenomenal*?

6 Interestingly, Fernández adds “falsely or not” at the end of this sentence. This suggests that he is not using the term “remembering” in its *factive* sense. Given the commitments of the theory, this is as it should be.

7 As Lewis (1966) points out in his classic treatment, “it is usually easy to find conditions which are *almost* necessary and sufficient for an experience” (p. 22). In relaxing the constitutive conditions for being in a mental state, functionalists employ the strategy popularized by analytic behaviorism. In any case, that such a relaxation is necessary when discussing phenomena studied by the (special) sciences is now a theoretical commonplace. See also note 8.

2 Relaxation and Realization

Let's start with a simple point. As welcome as the relaxation of constitutive conditions may be, it cannot be unprincipled. A theory which tells us that "memories are caused by past experiences *except when they aren't*" would ring a lot of alarm bells and be rightfully treated as explanatorily suspect.⁸ Functionalists, thus, may not be able to provide a full catalogue of exceptions, but they do owe us an explanation as to *why* such exceptions occur. Two main, and closely connected, reasons can be found in the literature. First, the world is complex and messy, so we shouldn't expect the regularities, psychological or otherwise, to be "tidy" (see, e.g., Pietroski & Rey, 1995). Second, mental states (at least in our world) are realized by physical systems, and –if the functionalists are to be trusted– can indeed be *multiply* realized by different (kinds of) physical systems. Given that external factors can sometimes affect the functioning of such systems, it would be unreasonable to expect that every token mental state will actually bring about its "proprietary" effects. The limit case is clear. A memory you entertain at t_1 will not necessarily cause a disposition to believe something at t_2 since, after all, you may be killed between t_1 and t_2 . This standard picture, which Fernández more or less inherits, raises all sorts of pesky questions, which have annoyed philosophers for some time (see, e.g., Fodor, 1991; Schiffer, 1991). One matters here: in virtue of what does a *token* mental state count as a realization of a functional kind (such as memory)?⁹

To see the issue, consider a candidate realizer state, which has the relevant representational profile, yet doesn't bear the causal relation we're interested in –i.e. it is *not* caused by a past experience. For example, suppose that I have a *seeming* recollection of celebrating a past birthday, which causes in me both a belief that this celebration happened and a belief that I've experienced it.¹⁰ As it happens, however, the state was not actually caused by my past experience. What should we make of such a token state? Is this a memory that simply plays the mnemonic role imperfectly? Or is it a realization of a different type of state, with a relevantly dissimilar functional profile (say: an imagining)?¹¹ Indeed, is there a principled way of deciding between these two verdicts? Attempting to show how FTM avoids the metaphysical stringency of causal theories, Fernández considers a similar class of cases –so-called "embellishment"

8 This issue is closely connected to the concern that so-called "hedged" (or *ceteris paribus*) laws are explanatorily vacuous. If they are vacuous, they *seemingly* cannot play a role in empirical science. And, given that the sciences (and not just the "special" ones) regularly posit such laws, this is a serious problem.

9 In what follows, I will use 'kind' and 'type' interchangeably.

10 So as not to complicate things, I assume that the mental state in question bears the *other* causal relation constitutive of memories (i.e. it does cause the relevant kinds of beliefs). As far as I can see, this doesn't affect the arguments below.

11 This is, of course, a live possibility. After all, there are token states –e.g. seemingly recollecting one's *second* birthday– which we strongly suspect to be (mere) imaginings in disguise (for a variety of relatively good reasons; see McCarroll, 2020). There is no reason to think that token states, which are *less obviously* realizations of imaginings, do not exist.

cases—, in which subjects seemingly recollect features of events they haven't previously experienced (2018, pp. 55-56; 2019, pp. 37-38). In a striking move, Fernández then appeals to his *intuitions* about the kindhood of these token states: “it *seems to me* that the mental [state] at issue *does* play the mnemonic role... [It] is the type of [state] that tends to be produced in me by past perceptual experiences” (2019, p. 51, *emphasis added*).¹² Taken on its own, this claim is quite puzzling. Fernández *does* have a story to tell about intuitions and their role in the specification of memories' functional roles (see 2019, Ch.1).¹³ It is, of course, subject to all sorts of methodological concerns, the generation of which has become a serious sport (Papineau, 2020). But whatever we think about the role of intuitions in characterizing functional kinds, we surely cannot expect to have intuitions about whether token states realize a kind. Whether a kind is actually realized depends on *how the world is*, and we have no reason to think that the intricate functional profiles of individual states would be transparent and available for introspection. Indeed, if we *had* reliable intuitions about the kindhood of token states, we wouldn't need to relax the constitutive conditions for remembering. We would just provide a more detailed account, which catalogs the exceptions from the relevant generalizations. Yet, we don't. In reality, we can't intuit whether a state functions as a memory any more than we can intuit whether a state functions as a resting state of a sodium channel.

So, how should we think about token states? When, and in virtue of what, do they count as realizations of functional kinds? The answer, already foreshadowed a few times, is straightforward: a token state realizes a functional kind just in case it is embedded in a *system* with the proper organization, specified by the relevant functionalist theory. Michael Antony (1994) articulates the idea well:

[T]he token must be properly situated in a system of the right sort... A system can be conceived of, roughly, as a set of physical conditions that allow for specific sorts of causal interactions among tokens. A system is of the right sort if the specific types of causal interactions it supports map appropriately onto the set of causal relations dictated by the functionalist theory in question. And a token is properly situated in the system for it to be an instance of a functional type F if it gets paired with F in the aforementioned mapping (in virtue of instantiating an appropriate first-order property) (p. 112, emphasis original).

In our case, a realizing system will be of the right sort if the kinds of causal interactions it supports map appropriately onto the set of causal relations specified by Fernández's

12 Fortunately, this intuition-supported claim doesn't *exhaust* Fernández's treatment of the issue. See below.

13 On Fernández's account, intuitions also play a role in determining the truth-conditions of episodic memories. Since his answers to the “metaphysical” and the “intentionality” questions are intended to be independent (see 2019, Ch.1), I bracket this issue here. For the limitations of intuition-driven *a priori* functionalism (of the kind Fernández endorses), see next section.

functionalist theory. Accordingly, a token state will be an instance of the functional type *memory* iff it is paired with it in such a mapping (The functional type is a *second-order* state, which is realized by this first-order token state but can, at least in principle, be realized by different first-order states; see Fernández, 2019, pp. 47-56). As long as the realizing system is of the right sort –and remains intact– a properly situated token state can realize the mnemonic role even if it *never* actually bears the constitutive causal relations (e.g., in a case when the organism is destroyed). So, whether our candidate state above is a memory depends on whether it is properly situated in a system of the right sort. And, of course, whether *this* is the case is by-and-large an empirical question (see next section). Fernández, indeed, appreciates the point, observing that, *as a matter of fact*, the “faculties of perception and memory are related in such a way that perceptual experiences [...] do produce [states that realize the mnemonic role]” (2019, p. 51). Yet, he needlessly runs it together with the intuition-supported claim discussed above.¹⁴

The account sketched above sits very well with a prominent strategy for characterizing *ceteris paribus* (CP) laws of the kind functionalists often traffic in (e.g., “other things being equal, memories cause beliefs of a specific kind”). On this strategy, formidably defended by Pietroski & Rey (1995), CP-laws hold only in systems considered in abstraction from external, independently existing factors. They allow for “abnormal” instances because such factors can, and indeed routinely do, affect the functioning of systems under consideration (the world being messy and all). Just how we should understand this notion of independence is one of those pesky questions, which –thank goodness– I’ll not examine here.¹⁵ It is worth noting, however, that we need not rely on an excessively robust metaphysics of systems in order to cash it out. We may, rather, think of the treatment of systems *in isolation* as part of the regular scientific practice of abstraction and idealization –i.e., ignoring some aspects of a phenomenon with the goal of understanding others (This is, indeed, the gloss offered by Pietroski & Rey, 1995). So, an ecumenical pluralism, or a Craver-style perspectivalism (2013), would in principle do just as well.

3 Causation and Mnemonic Roles

Back to business. Here’s the rub: the debate between modern causalists and their foes is best understood as a debate about the *existence* (in human beings) of systems which

¹⁴ Thus, Fernández finishes the sentence cited above by concluding that a candidate *token* state he is entertaining is an instantiation of the relevant functional kind (memory). As I’ve argued in the main text, it’s not clear why he thinks that conclusion is warranted.

¹⁵ Pietroski & Rey (1995) defend a specific conception of independence, on which factors whose *only* role is to save the proposed CP-laws (in an *ad hoc* manner) are excluded.

support kinds of causal interactions that map onto the set of relations dictated by causal theories of memory. In other words, it is about whether the faculties of memory and perception, roughly speaking, are *actually* related in the way causal theorists have traditionally supposed they are. So, short of begging the decisive empirical question, Fernández's functionalism cannot help us settle *this debate*.

Consider the simulationist position first. Michaelian is quite clear that this is how he understands the dialectic. On the framing offered in his *Mental Time Travel* (2016a), the empirical results—which, among other things, show a close processing connection between memory and (future-oriented) imagination—are taken to be “surprising” *precisely* because they upend the picture that has dominated the conceptualization of memory since at least the 1960s.¹⁶ While Michaelian's general attitude is clear, he sometimes wavers between two versions of his view (in a way, indeed, that may make some of his readers uncomfortable). On the *weak* version, he takes the emerging empirical evidence to present a challenge to causal theories, which philosophers should take really seriously.¹⁷ This claim is often paired with a programmatic statement about the weight we should put on such evidence when doing philosophy of mind (e.g., Michaelian, 2016b, pp. 65–67). On the *strong* version, responsible for much of the hype and infamy of the view, the evidence does more than present a serious challenge. It actually shows causal theories to be false. Consider the way in which the empirical picture is leveraged to provide an argument against causalism:

Since imagining a future event trivially does not presuppose the existence of a causal connection between the subject's thought of the event and his experience of the event, this, in turn, suggests that remembering a past event likewise should not be taken to presuppose the existence of a causal connection between the subject's thought of the event and his experience of the event: the episodic memory system [...] appears simply not to be designed in such a way that the presence of such a connection can be taken for granted in every case of genuine remembering (Michaelian & Sant'Anna, 2019, p.14, *emphasis added*).

In other words, the system that realizes the memory kind (in human beings) *actually* turns out not to instantiate the functional organization specified by causal theories.

16 His psychological counterparts are equally clear about their commitments. Indeed, they frequently characterize the emerging empirical picture as “revolutionary” or “paradigm-shifting”. See, e.g., Schacter (2008, p. 5): “[T]he study of memory has undergone dramatic changes during the past couple of decades, some even revolutionary [...] We now know enough to demolish [a] long-standing myth: that memories are passive or literal recordings of reality”. Schacter then goes on to catalog the number of different ways in which scientific developments challenge traditional conceptions of memory.

17 See, e.g., Michaelian (2016b): “The picture that emerges of a fully symmetrical ability to mentally travel backward and forward in time suggests that *philosophers would do well to follow the lead of psychologists*, attending more closely than they have historically done to our capacity for future-directed episodic thought and to the ways in which that capacity gives rise to knowledge of future events” (p. 63, *emphasis added*). For the importance of this “symmetry” for our purposes, see the strong version of Michaelian's view.

Memory, that is, is not related to perception in quite the way traditional theorists have supposed. Now, there may be cases –perhaps indefinitely many– in which a memory will be causally connected to a past experience. Yet, this should not obscure the purported fact that memory systems are deeply pragmatic, as it were, routinely utilizing information from a variety of sources (more on this below). While there may be good reasons to resist this simulationist conclusion, we should take the claims at face value.

Causal theorists have indeed attempted to resist the conclusion in a variety of ways. Robins, for example, has done important work trying to demonstrate that the strong simulationist thesis is not warranted. Thus, in Robins (2020a), she argues that endorsing a dynamic view of memory processing doesn't necessitate the abandonment of preservationist theories.¹⁸ In Robins (2020b), she takes on Michaelian directly, arguing that the empirical evidence does *not* license the conclusion that memory and (future-oriented) imagination are states/processes of the same kind (note the role the kindhood claim plays in Michaelian's argument against causalism presented above). To put her points in our idiom, there may *be* realizer systems supporting the kinds of causal interactions that map onto the relations specified by a –suitably amended, to be sure– *causal* theory.¹⁹ Werning (2020) reaches a similar conclusion about the necessity of a causal connection in remembering, but his amendment of traditional causal theories is more drastic. On his view, the causal connection to a past experience does important “work”, securing the reliable production of accurate representations by memory systems. Yet, it is not sustained by *content*-bearing memory traces (in an important sense, then, Werning's “minimalist” theory is causal but *not* representational). While his proposal is speculative, it resonates well with an exciting and fruitful research program in the neurosciences (predictive processing). More importantly for our purposes, it illustrates clearly that nothing in the available empirical evidence compels us to accept the simulationist claim about memory causation. Indeed, according to Werning, we are compelled to reject it.²⁰

18 Robins (2020a) focuses on *neural* dynamics, arguing for its compatibility with the existence of an engram, which she characterizes as “a refashioning of the age-old memory trace: the entity responsible for forming, storing, and retrieving memories” (p. 1131). How faithful this refashioning actually is a matter for another occasion.

19 Here I continue to use “system” in the way defined (by Antony) in section 2. Note that this notion is “thinner” than the one offered by Michaelian, who takes systems to be to functionally individuated (sets of) mechanisms, individuated in computational, representational and neural terms (see his 2016a, Ch. 2).

20 A worry a reader may have here, which an anonymous referee does have, is that these views are too dissimilar to be grouped under the umbrella of “causal” theories. I am worried about this myself (indeed, this paper aims to illustrate that causal theories may be developed in some surprising ways). Yet, for the purpose of providing a *provisional sketch of the disagreement with STM*, causal theories may be characterized as those theories that endorse what Michaelian & Robins (2018, p. 24) call the “core claim”: that a memory has to be actually caused by a past experience. That said, it is interesting to consider whether, e.g., Werning's “minimal” traces can sustain an *appropriate* causal connection between memories and past experiences. If they can't, then the minimalist causal theory will indeed be quite different from traditional causal approaches. I am thankful to a referee for prompting me to add this clarification.

At this point, it should be obvious why Fernández's functionalism doesn't move the needle in the debate between causalists and simulationists. The debate concerns the instantiation of systems of the right sort, and the truth about such instantiation is, as Scully and Mulder used to say, *out there*. Now, the functionalist may insist that theories of the kind Fernández offers provide important constraints about the kinds of states we *can* discover when searching for memories. And, perhaps they are right about this.²¹ Yet, two things should be kept in mind. First, the existence of constraints à la Fernández does not entail that a "stronger" theory—along causalist lines—will not end up being vindicated.²² On the flip side, if it turns out that the kinds of relations posited by FTM are *not* regularly instantiated, then reserving the name "memory" for a non-actualized kind would be a poor consolation.

What, then, is the relation between FTM and the theories introduced above? Consider Michaelian's simulation theory (STM) first. FTM and STM are sometimes grouped together as "postcausal" on the grounds that the two theories reject "the core claim of the causal theory": that a memory has to be *actually* caused by a past experience (e.g., Michaelian & Robins, 2018, p. 24). In fact, there have been recent attempts to combine them into one—I am tempted to call it a "superfunctionalist"—theory (Langland-Hassan, 2021). I hope that the discussion above hints at why we should be very careful when endeavoring to do so. While it is indeed the case that, *in a sense*, both FTM and STM "relax" the causal condition, they do so for very different reasons. If my analysis is correct, Fernández's functionalist should motivate the relaxation by pointing to the fact that external factors will sometimes/often interfere with the functioning of memory systems (the world *is* messy). When we specify the causal relations *constitutive* of the functional role of memories, however, we should abstract away from such interference. In other words, theories of memory should consider the functioning of memory systems in "ideal" circumstances.²³ And, in ideal circumstances, memories will presumably *always* bear their constitutive causal relations. Things look quite different on STM. For the simulationist, it is *not* simply the case that memories occasionally do not bear causal relations that are nevertheless constitutive of their functional roles. Rather, a causal connection to a past experience is not constitutive of memories' functional roles, even if does in fact sometimes/often obtain. On the

21 To put the point simply, if we find mental states that have none of the features we typically associate with memories, then we have a relatively decent reason to think that they are *not* in fact memories. That said, what we should do when we find mental states that have *some* of these features, but not others, is a million-dollar question.

22 Note that causal theories are stronger only relative to the constitutive relation this essay focuses on (i.e., the causal connection between memories and past experiences). Many causal theories are weaker than FTM relative to the other constitutive relation posited by the theory. That is, they do not take the formation of specific beliefs to be necessary for (or even typical in) remembering.

23 Idealization of this kind is arguably omnipresent in scientific theorizing. For example, on a prominent view of psycholinguistics, the discipline is concerned with an idealized linguistic *competence*, unaffected by external conditions on linguistic *performance*, such as memory limitations or shifts of attention (Chomsky, 1965, pp. 1-8).

theory, a subject remembers a past event just in case they entertain a representation produced a properly functioning episodic system, which ‘aims’ to represent an event from the subject’s personal past (Michaelian, 2016a, p. 107). What is constitutive of the functional role of memories, then, is that they are produced by such a system and such an operation. Michaelian insists that even when the ‘aim’ is to produce an event from the subject’s personal past, the system *need not* draw on information originating in the past experience of the event. This is despite the fact it may often be most “efficient” to draw on some such information (Michaelian, 2016a, p. 104).²⁴ It is important to realize that, unlike in the case of FTM, this is a claim about the functioning of the episodic system in *ideal* circumstances. Even if we abstract away from all external interference, it will not be the case that all mnemonic representations will bear causal connections to past experiences. Hence, strictly speaking, Michaelian doesn’t offer a relaxation but a full-blown *rejection* of the causal condition for remembering. In this sense, the functional organization of memory systems, as described by STM, is quite unlike the organization specified by causal or epistemic theories. Michaelian offers a number of reasons for this proposal, which of course I can’t properly assess here. Yet, the key claim seems to be the one sketched above. Since the purported episodic memory system turns out to be employed in a variety of activities other than remembering—e.g. in future-oriented or counterfactual imagining—its operations will understandably be quite different from what traditional theories have expected them to be. Given all of this, then, we should be wary of grouping FTM and STM under the general umbrella of “postcausal” theories. Moreover, the simulationist would also be troubled by the aprioristic flavor of FTM, a flavor they have tried really hard to get rid of.

So, what about causal theories? FTM may not really be postcausal in the same sense STM is, but it surely *is* postcausal, right? (After all, it does *seemingly* reject the core claim of the causal theory). If I am right about the landscape of theses motivating functionalism, then the answer is: not necessarily. We may indeed be able to formulate a causal theory in a way that is congenial to the spirit, if not the letter, of our functionalist. A causalist may argue—in quite good company (Boyd, 1991; Fodor, 1974; Pöyhönen, 2015)—that psychological kinds like remembering are unlikely to participate in “strict” natural laws. Yet, they may still insist that, in an important sense, causal theories are right about the necessity of a causal condition in remembering. Hence, they may choose to formulate the key claim of their theory as a CP-law: “*Other things being equal*, all memories are appropriately causally connected to past experiences”. Now, the thing to notice is that, given that this nomic generalization will include a CP-clause, the allowed abnormal instances (in which memories will *not* be causally connected to past

24 In just how many cases information from a past experience is actually used is, of course, an empirical question. What is thus worth highlighting is that, in the absence of good evidence, Michaelian’s *a priori* argument from efficiency should be taken with some salt.

experiences) will *not* constitute exceptions to it. The generalization, i.e., will not purport to hold in *all* conditions since external factors can, and sometimes do, interfere with the functioning of memory systems. Yet, in the conditions in which it does hold –the ideal conditions appropriate for theory building– it *will* purport to be exceptionless.²⁵ (The causal theorist will thus be committed to the claim that there won't be any abnormal cases that cannot *in principle* be explained by the occurrence of external factors). This impressionistic sketch leaves many important and difficult questions open. Yet, a “hedged” theory of this kind may nevertheless be attractive to some causalists.²⁶ Two key points are worth flagging here. First, to show that such a theory is false, it will be not enough to show that not all memories are causally connected to past experiences, as is sometimes (uncritically) assumed. What needs to be shown is rather that these are not *abnormal* cases of remembering –a much taller order. Second, and important for our purposes, a hedged causal theory will end up looking surprisingly similar to FTM.²⁷ Pace Michaelian & Robins (2018), Fernández's functionalism may not be genuinely post-causal after all. The devil is in the details.

The general lesson of this section should be familiar to functionalists concerned about the limits of Lewis-style analysis. As Rey (1997, p. 187) puts it in a classic presentation: “At best, what would seem to be available from *a priori* analysis would be some rough constraints that merely ‘fix the reference’ of mental terms... But to determine what kind [they pick out], we would need the aid of an *empirical* psychological theory”. The constraints provided by Fernández's FTM may indeed be suitable (if rough) for fixing the reference of “remembering”. Yet, to uncover the intricate functional profiles of states of remembering, we will need to consult a more detailed, and fully fleshed out, psychological theory.

4. Conclusion

Fernández's (2018; 2019) functionalist theory of memory aims to relax the causal condition for remembering. On the theory, token memories need not bear causal connections to past experiences. They only have to play the mnemonic functional role, which is constituted by a causal connection to the past. In this paper, I argued that

25 Cf. Pietroski & Rey (1995, p. 88): “Let us say that a ‘strict’ law is one that contains no cp-clause, even implicitly [...] One might hope that some future unified field theory will provide an example of a strict law. We do claim that not all laws are strict in this sense. Indeed, given current science, the appropriate question would seem to be whether any laws are strict. But even if they are not strict, still they may be (and we grant that they ought to be) exceptionless. So, if there are exceptions to [a purported law] then *a fortiori* [it] is not a law”.

26 Werning (2020), if I read him correctly, has some sympathy for such a view (see, e.g., footnote 12). Robins (personal communication) seems less sympathetic to the proposal.

27 At least relative to the *backward-looking* causal profile of memories. Causalists may have different views about the tendency of memories to cause relevant beliefs.

this relaxation doesn't advance the debate between causal and simulation theories. A token mental state can be taken to play a mnemonic role only if it is embedded in a system of the right sort and the debate concerns precisely the existence of such systems. Moreover, if we examine the most plausible functionalist motivation for relaxing the causal condition—the presence of external interference on memory systems—we may learn something surprising about the theoretical landscape. Fernández's functionalist theory is not postcausal in the same sense the simulation theory is and may indeed not be postcausal at all.

References

- Andonovski, N. (2020). Singularism about episodic memory. *Review of Philosophy and Psychology*, 11, 335–365. <https://doi.org/10.1007/s13164-020-00464-y>
- Andonovski, N. (2021). Memory as triage: facing up to the hard question of memory. *Review of Philosophy and Psychology*, 12, 227–256. <https://doi.org/10.1007/s13164-020-00514-5>
- Antony, M. V. (1994). Against functionalist theories of consciousness. *Mind & Language*, 9(2), 105–123. <https://doi.org/10.1111/j.1468-0017.1994.tb00219.x>
- Block, N. (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science*, 9, 261–325
- Boyd, R. N. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61, 127–148. <https://doi.org/10.1007/BF00385837>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Craver, C. F. (2013). Functions and mechanisms: a perspectivalist view. In P. Huneman (Ed.), *Functions: selection and mechanisms* (pp. 133–158). Springer. https://doi.org/10.1007/978-94-007-5304-4_8
- Fernández, J. (2018). The functional character of memory. In K. Michaelian, D. Debus & D. Perrin, (Eds.), *New directions in the philosophy of memory* (pp. 52–72). Routledge. <https://doi.org/10.4324/9781315159591-4>
- Fernández, J. (2019). *Memory: a self-referential account*. Oxford University Press. <https://doi.org/10.1093/oso/9780190073008.001.0001>
- Fodor, J. A. (1974). Special sciences or the disunity of the sciences as a working hypothesis. *Synthese*, 28, 97–115. <https://doi.org/10.1007/BF00485230>
- Fodor, J. A. (1991). You can fool some of the people all of the time, everything else being equal; hedged laws and psychological explanations. *Mind*, 100(397), 19–34. <https://doi.org/10.1093/mind/C.397.19>
- Kim, J. (1993). *Supervenience and mind: selected philosophical essays*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511625220>

- Langland-Hassan, P. (2021). What sort of imagining might remembering be? *Journal of the American Philosophical Association*, 1-21. <https://doi.org/10.1017/apa.2020.28>
- Lewis, D. K. (1966). An argument for the identity theory. *The Journal of Philosophy*, 63(1), 17-25. <https://doi.org/10.2307/2024524>
- Martin, C. B. & Deutscher, M. (1966). Remembering. *Philosophical Review*, 75(2): 161-196. <https://doi.org/10.2307/2183082>
- McCarroll, C. J. (2020). Remembering the personal past: beyond the boundaries of imagination. *Frontiers in Psychology*, 11, 2652. <https://doi.org/10.3389/fpsyg.2020.585352>
- Michaelian, K. (2016a). *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. (2016b). Against discontinuism: mental time travel and our knowledge of past and future events. In K. Michaelian, S. B. Klein & K. Szpunar (Eds.), *Seeing the future: theoretical perspectives on future-oriented mental time travel*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190241537.001.0001>
- Michaelian, K. & Robins, S. (2018). Beyond the causal theory? Fifty years after Martin and Deutscher. In: K. Michaelian, D. Debus & D. Perrin (Eds.), *New directions in the philosophy of memory* (pp. 13-32). Routledge. <https://doi.org/10.4324/9781315159591-2>
- Michaelian, K. & Sant'Anna, A. (2019). Memory without content? Radical enactivism and (post) causal theories of memory. *Synthese*, 1-29. <https://doi.org/10.1007/s11229-019-02119-7>
- Papineau, D. (2020). Naturalism. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/naturalism/>
- Pietroski, P. & Rey, G. (1995). When other things aren't equal: saving ceteris paribus laws from vacuity. *The British Journal for the Philosophy of Science*, 46(1), 81-110. <https://doi.org/10.1093/bjps/46.1.81>
- Pöyhönen, S. (2015). Memory as a cognitive kind: brains, remembering dyads and exograms. In C. Kendig (Ed.), *Natural kinds and classification in scientific practice* (pp. 145-156). Pickering & Chatto.
- Rey, G. (1997). *Contemporary philosophy of mind: a contentiously classical approach*. Wiley-Blackwell
- Robins, S. K. (2016). *Representing the past: memory traces and the causal theory of memory*. *Philosophical Studies*, 173, 2993-3013. <https://doi.org/10.1007/s11098-016-0647-x>
- Robins, S. K. (2020a). Stable engrams and neural dynamics. *Philosophy of Science*, 87(5), 1130-1139. <https://doi.org/10.1086/710624>
- Robins, S. K. (2020b). Defending discontinuism. *Naturally. Review of Philosophy and Psychology*, 11, 469-486. <https://doi.org/10.1007/s13164-020-00462-0>

- Rupert, R. D. (2006). Functionalism, mental causation, and the problem of metaphysically necessary effects. *Noûs*, 40(2), 256-283. <https://doi.org/10.1111/j.0029-4624.2006.00609.x>
- Schacter, D. L. (2008). *Searching for memory: the brain, the mind, and the past*. Basic Books
- Schiffer, S. (1991). *Ceteris paribus laws*. *Mind*, 100(397), 1-17. <https://doi.org/10.1093/mind/C.397.1>
- Werning, M. (2020). Predicting the past from minimal traces: episodic memory and its distinction from imagination and preservation. *Review of Philosophy and Psychology*, 11, 301-333. <https://doi.org/10.1007/s13164-020-00471-z>